DURATION MODELS

Lecture Notes 2015-2016

Part 1

Introduction

Frederic PLANCHET / Olivier DURAND

Version 4.0

February 2016

Table of content

1. Introduction

1.1. Historical landmarks[1]

The formalised analysis of duration data goes back to the English school of political arithmetic – noticeably the work of John Graunt (1620-1674) and William Petty (1623-1687) at the time of the first studies on mortality in England at the 17th century (see Le Bras [2000]). The concepts of life expectancy and residual life expectancy are then defined.

The search for underlying laws for these phenomena starts at the 19th century – in particular with the formula suggested by Benjamin Gompertz in 1825 to model the probability of dying at age *x*:

$$h(x) = a \times b^x$$

This model (which is in fact a geometric progression of the rates of death with common ratio *b*) will be upgraded by William Makeham in 1860:

$$h(x) = c + a \times b^x$$

Survival durations will long remain a subject studied by demographers as well as actuaries, until the appearance of the theory of "reliability" for physical systems. In 1951 W. Weibull publishes in a journal of mechanics an article where he proposes the following form for the hazard function:

$$h(t) = \lambda \alpha t^{\alpha - 1}$$

Weibull's article noticeably deals with one of the important characteristics of duration data: the presence of truncated or censored data.

Two other important dates must be quoted. The 1958 article of E. Kaplan and P. Meier in which they propose to use in the medical field a nonparametric estimator making it possible to integrate the data censored introduced in 1912 by P. Böhmer, the "PL" estimator of the survival function.

In 1972 David Cox publishes an article providing the bases of an important particular case of "proportional hazard" model using (exogenous) determinants:

$$h(x) = e^{\beta z} h_0(x)$$

with β a vector of parameters (unknown) and h_0 the unknown base hazard function; it is therefore a semi-parametric model. This reference model led to many further developments and alternatives: introduction of a temporal evolution, taking into account of dependency between observed variables, stratification of the effect of covariables, *etc.*

Finally, in closing this brief panorama, one can mention two recent evolutions of duration models:

- ✓ Prospective tables and the two-dimensional models "age x year", the founding reference being Lee and Carter [1992].

[1] Largely inspired from Droesbeke *et al.* [1989].

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

- ✓ Quantification of the "non-poolable" part of mortality risk, *via* mortality stochastic models (see Cairns *et al*. [2004]).

1.2. Characteristics of duration data

The first characteristic of the duration data is to be generated by positive random variables; even though one can imagine restraining any real random variable to $[0,+\infty[$ through a well-chosen transformation (e.g. the exponential function), as a result of this characteristic the normal distribution cannot be the duration models' reference law.

Looking at random variables in terms of duration will also lead to a shift from using the cumulative distribution function towards using the survival function and the hazard function.

The third characteristic is the fact that the reference situation is that of incomplete data. This can be the consequence of:

- ✓ The fact that the random variable is observable only on a subset of $[0,+\infty[$; the model is then said to be truncated.
- ✓ The fact that for certain individuals the result of the experiment is only partially observed: for example, the experiment is of limited time T and for the individuals still alive in T, the lifetime will not be known – only that it is higher than T; the model is then said to be censored (or right-censored).

Lastly, duration data generally use exogenous determinants: for example, life expectancy depends on gender, socio-professional level, region, *etc.*

1.3. Statistical models

The various usual statistical models are found in the description of duration data:

- ✓ Parametric models: e.g. Makeham model.
- ✓ Nonparametric models: e.g. Kaplan-Meier estimator.
- ✓ Semi-parametric models: e.g. Cox model.

Stochastic models can also be added to this typology – they hold a special place ("layered" on one of the above models).

2. Representation of a survival distribution

One considers a random variable T with values in $[0,+\infty[$, and one denotes $F(t)=P(T \le t)$ its cumulative distribution function (continuous to the right). When the probability density function of T exists, it will be noted $f(t)=\frac{d}{dt}F(t)=\lim_{h\to 0}\frac{P(t \le T \le T+h)}{h}$.

2.1. Survival function

The survival function is the complementary cumulative distribution function:

$$S(t)=1-F(t)=P(T>t)$$

S is thus a decreasing function such that $S(0)=1$ (if $P(T=0)=0$, what we suppose) and $\lim_{t\to\infty} S(t)=0$. If the duration of survival expectancy exists, then it is expressed simply using S:

$$E(T)=\int_0^\infty t dF(t) = -\int_0^\infty t dS(t) = \int_0^\infty S(t)dt$$

<u>Demonstration</u>: It is supposed that the expectancy exists. Because $\int_0^\infty t dF(t) = \lim_{u\to+\infty}\int_0^u t dF(t)$; integrating by parts, one can write $\int_0^u t dF(t) = -\int_0^u t dS(t) = -uS(u)+\int_0^u S(t)dt$; the Markov inequality then ensures that $tS(t)\le E(T)$ and thus the term $uS(u)$ is bounded. One deduces that the integral $\int_0^\infty S(t)dt$ converges, which implies that $\lim_{t\to\infty} tS(t)=0$ and, pushing to the limit, one gets the expected result.

One can also demonstrate this result in the following way by observing that $\int_0^\infty S(t)dt = \int_0^\infty E\left(1_{\{T>t\}}\right)dt$ and by Fubini $\int_0^\infty E\left(1_{\{T>t\}}\right)dt = E\left(\int_0^\infty 1_{\{T>t\}}dt\right) = E\left(\int_0^T dt\right) = E(T)$.

In the same manner, one demonstrates that:

$$V(T) = 2\int_0^\infty tS(t)dt - E(T)^2 .$$

2.2. Conditional survival function

The conditional survival function is defined by $S_u(t) = P(T>u+t|T>u)$; one is interested in the survival of an element after time $u+t$, knowing that it already "functioned" properly until time u. Returning to the definition of conditional probability, one can write:

$$S_u(t) = P(T>u+t|T>u) = \frac{P(T>t+u)}{P(T>u)} = \frac{S(u+t)}{S(u)}.$$

The conditional survival function is thus expressed simply using the survival function.

2.3. Hazard function

The hazard function (or failure rate, rate of death, instantaneous risk, *etc*.) is by definition:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d}{dt}\ln S(t).$$

It follows that the hazard function entirely determines the law of T and that the following relation is obtained:

$$S(t) = \exp\left(-\int_0^t h(s)ds\right).$$

One generally notes $H(t) = \int_0^t h(s)\,ds$ as the "cumulated hazard function", which is such that $S(t) = \mathbf{exp}(-H(t))$ and which is obviously increasing. The fact that *H(T)* follows an exponential law of parameter 1 is used in some goodness-of-fit tests. This property results from:

$$P(H(T) > x) = P(T > H^{-1}(x)) = S(H^{-1}(x)) = \mathbf{exp}(-H(H^{-1}(x))) = \mathbf{exp}(-x).$$

Combining this with the definition of the conditional survival function, one obtains:

$$S_u(t) = \mathbf{exp}\left(-\int_u^{u+t} h(s)\,ds\right).$$

In other words, the hazard function of survival conditional to the fact of being "functional" at time *u* is simply $t \to h(u+t)$. One deduces that the hazard function is increasing if and only if the residual lifetime after *u* is stochastically decreasing[2] as a function of *u*.

It is often the hazard function which is used to specify a duration model. It indeed offers a "physical" interpretation; by using the definition of both hazard function and survival function one can write:

$$h(t) = \lim_{u \to 0} \frac{P(t < T \le t+u)}{uS(t)} = \lim_{u \to 0} \frac{P(t < T \le t+u \mid T > t)}{u}$$

which means that for "small" values of *u*, *h(t)u* is roughly the probability that the component breaks down between times *t* and *t+u*, knowing that it is "functional" in *t*. In other words:

$$P(t < T \le T + dt \mid T > t) = h(t)\,dt.$$

2.4. The case of discrete variables

If the random variable *T* takes integer values, its distribution is described by $p_k = P(T = k)$, for $k \ge 0$. The survival function is simply written $S(k) = \sum_{m \ge k+1} p_m$.

The interpretation of the hazard function given in 2.3 above naturally leads to the discrete case as follows:

$$h(k) = P(T = k \mid T > k-1) = \frac{p_k}{S(k-1)}.$$

The hazard function at the point *k* is thus interpreted as the rate of death at the age *k*. Therefore $1 - h(k) = \frac{S(k)}{S(k-1)}$, and then, by recurrence:

$$S(k) = \prod_{m=1}^{k} (1 - h(m)).$$

[2] By definition *X* is stochastically larger than *Y* if $S_X(t) \ge S_Y(t)$.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

The survival function is thus assimilated to the (L_x) of a mortality table, while the hazard function is assimilated to its (q_x). It will be noted that any continuous representation of a lifetime T can be associated to a discrete representation, using $X = k, k \leq T < k+1$ (in other words $X = [T]$).

In practice however, the problems are in general of opposite nature: a discrete distribution is estimated and then rates of death are wanted at any age. For that purpose a hypothesis has to be made, which allows passing from a discrete to a continuous expression of the law; three approaches are classically used:

- ✓ Linear transformation of the survival function, which is similar to assuming uniform distribution of the exits on $[k,k+1[$ (UDD assumption);
- ✓ The assumption of permanence of the hazard function on $[k,k+1[$, which leads to an exponential form;
- ✓ The "Balducci assumption", which leads to a hyperbolic form.

These three approaches are summarised in the table below[3]:

Fonction	Linéaire (DUD)	Exponentielle (Force constante)	Hyperbolique (Balducci)
l_{x+t}	$l_x - t \cdot d_x$	$l_x \cdot (p_x)^t = (l_{x+1})^t \cdot (l_x)^{1-t}$	$\left[(1-t)\cdot\frac{1}{l_x} + t \cdot \frac{1}{l_{x+1}}\right]^{-1}$
${}_t p_x$	$1 - t \cdot q_x$	$(p_x)^t = e^{-\mu \cdot t}$	$\frac{px}{1-(1-t)\cdot q_x}$
${}_t q_x$	$t \cdot q_x$	$1-(1-q_x)^t$	$\frac{t \cdot q_x}{1-(1-t)\cdot q_x}$
${}_{1-t} p_{x+t}$	$\frac{p_x}{1-t\cdot q_x}$	$(p_x)^{1-t} = e^{-\mu \cdot (1-t)}$	$1-(1-t)\cdot q_x$
${}_{1-t} q_{x+t}$	$\frac{(1-t)\cdot q_x}{1-t\cdot q_x}$	$1-(1-q_x)^{1-t}$	$(1-t)\cdot q_x$
μ_{x+t}	$\frac{q_x}{1-t\cdot q_x}$	$\mu = -\ln p_x$	$\frac{q_x}{1-(1-t)\cdot q_x}$
${}_t p_x \cdot \mu_{x+t}$	q_x	$\mu \cdot e^{-\mu \cdot t}$	$\frac{q_x \cdot (1-q_x)}{[1-(1-t)\cdot q_x]^2}$
L_x	$l_x - \frac{1}{2}\cdot d_x = l_{x+1} + \frac{1}{2}\cdot d_x$	$\frac{d_x}{\mu}$	$-l_{x+1}\cdot\frac{\ln p_x}{q_x}$
m_x	$\frac{q_x}{1-\frac{1}{2}\cdot q_x}$	μ	$\frac{(q_x)^2}{-p_x \cdot \ln p_x}$

In the rest of this lecture the exponential form will generally be used (constant force). In some particular cases, noticeably parametric, it is not necessary to formulate an assumption – the model imposing the form to be used (see Makeham model for example).

[3] Table extracted from Langmeier [2000].

3. Usual parametric distributions

Only the most widely used models are detailed hereafter; as a general rule, all distributions used to model positive variables (lognormal, Pareto, logistic, *etc.*) can be used in duration models[4]. However, the basic distribution of parametric duration models is the exponential distribution – and its various generalisations – for reasons which will be developed *infra*[5].

The choice of the model determines the form of the hazard function; models with monotonic hazard function will be considered distinctly of those displaying "bell-shaped" or "U-shaped" hazard functions; these last models are seldom used in insurance, the reference situation being a hazard rate increasing (in the broad sense) with time.

3.1. Exponential model

The simplest specification consists in stating $h(t) = \lambda$, with $\lambda > 0$. Therefore $S(t) = e^{-\lambda t}$.

The exponential model is characterised by the fact that the conditional survival functions $\{S_u(.), u > 0\}$ are exponential with identical parameter $\lambda > 0$. Which means that the behaviour of the random variable T after time u is unrelated to what occurred up to that moment. It is also characterised by the fact that its survival function is multiplicative: $S(u+t) = S(u)S(t)$. These properties result from the expression of the conditional survival function presented in 2.2 above.

One can verify through direct calculation that $E(T) = \frac{1}{\lambda}$ and $V(T) = \frac{1}{\lambda^2}$. Estimating the parameter λ is standard procedure, starting from $L(\lambda) = \lambda^n \mathbf{exp}\left(-\lambda \sum_{i=1}^{n} T_i\right)$ which leads to $\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} T_i} = \frac{1}{\overline{T}}$.

3.2. Weibull model

It is supposed here that the hazard function is of the following form:

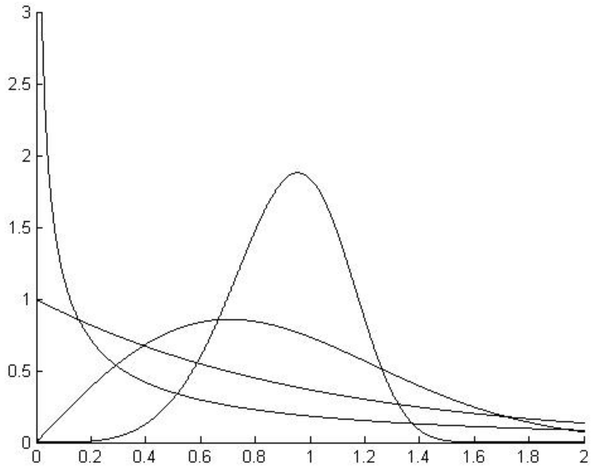
$$h(t) = \lambda \alpha t^{\alpha-1},\ \alpha, \lambda > 0$$

λ is a scale parameter and α a form parameter. This is a simple generalisation of the exponential model, making it possible to obtain increasing hazard functions with t if $\alpha > 1$ (then we have "wear") and decreasing with t if $\alpha < 1$ (then we have "break-in"). When $\alpha = 2$ and $\lambda = \frac{1}{2}$ this model is called "Rayleigh model"; it is used in physics to model the lifetime of particles or the noise at exit of transmission receivers[6].

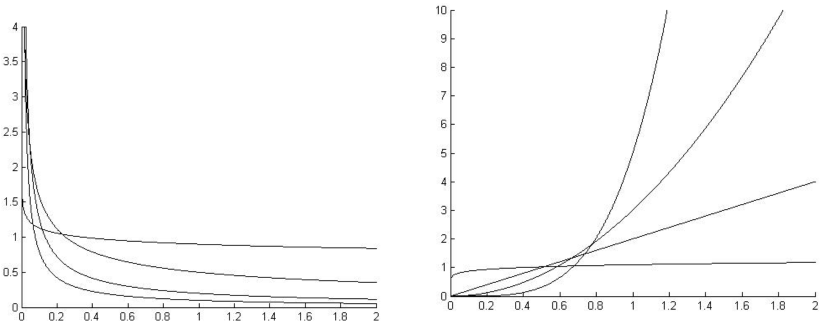
---

[4] For the properties of the usual distributions, see for example Partrat and Besson [2004].
[5] See the lecture notes on "Poissonnian Processes and queues"
[6] The Rayleigh distribution is also that of a reduced centered Gaussian vector.

The distribution of T is then the Weibull distribution, $W(\alpha,\lambda)$ whose survival function is written $S(t) = e^{-\lambda t^\alpha}$, $t > 0$. One can in particular notice that if the variable T is distributed according to an exponential law of parameter $\lambda > 0$ then $T^{1/\alpha}$ follows $W(\alpha,\lambda)$[7]. Depending on $\alpha > 0$ one can obtain very different forms of the probability density function[8]:



The hazard function is monotonic and looks like this:



$\alpha < 1$ $\alpha > 1$

The moments are obtained by observing that $E(T^k) = \lambda^{-k/\alpha}\Gamma\left(\frac{k}{\alpha}+1\right)$ with $\Gamma(x) = \int_0^{+\infty} u^{x-1}e^{-u}du$. To demonstrate this equality one writes the density of the Weibull distribution using $f(t) = S(t)h(t)$, which gives $f(t) = \lambda\alpha t^{\alpha-1}\mathbf{exp}(-\lambda t^\alpha)$. Therefore:

$$E(T^k) = \int_0^{+\infty} \lambda\alpha t^{k+\alpha-1}\mathbf{exp}(-\lambda t^\alpha)dt\,.$$

The change of variable $u = \lambda t^\alpha$ allows moving forward and yields:

[7] This gives a simple method to simulate realisations of the Weibull random variable.
[8] The graph is built with $\lambda = 1$.

$$E(T) = \lambda^{-1/\alpha}\Gamma\left(\frac{\alpha+1}{\alpha}\right) \text{ and } V(T) = \lambda^{-2/\alpha}\left(\Gamma\left(\frac{2}{\alpha}+1\right) - \Gamma^2\left(\frac{1}{\alpha}+1\right)\right)$$

The expression of the variance is the direct consequence of $V(T) = E(T^2) - E(T)^2$. One deduces from these expressions a remarkable property of the Weibull distribution, which is that the coefficient of variation $\frac{\sigma(T)}{E(T)}$ does not depend on the scale factor λ.

If $X = \ln(T)$ then $P(X \le x) = P(T \le e^x) = 1 - \exp(-\lambda e^{\alpha x})$, which can be written as follows:

$$P(X \le x) = 1 - \left(\exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right)\right)$$

while posing $\mu = \frac{1}{\alpha}\ln\left(1/\lambda\right)$ and $\sigma = 1/\alpha$. This is the Gumbel distribution (or double exponential), which is one of the three possible laws for the asymptotic distribution of the maximum of an iid sample[9].

The Weibull distribution appears naturally in the study of the asymptotic distribution of the minimum of an iid sample. Indeed, if $(X_1, \ldots, X_n)$ is a sample of a random variable of function of distribution G on $]0,+\infty[$ which behaviour near the origin verifies:

$$\lim_{x \to 0^+} \frac{G(x)}{\lambda x^\alpha} = 1$$

then $n^{1/\alpha} X_{(1)}$ converges in law, when n tends towards infinity, towards a distribution $W(\alpha, \lambda)$.

<u>Demonstration</u>: One has $P\left(n^{1/\alpha} X_{(1)} > x\right) = \left[1 - G\left(\frac{x}{n^{1/\alpha}}\right)\right]^n$ and thus:

$$\ln\left(P\left(n^{1/\alpha} X_{(1)} > x\right)\right) = n \ln\left[1 - G\left(\frac{x}{n^{1/\alpha}}\right)\right]$$

$$= n \ln\left[1 - \lambda\left(\frac{x}{n^{1/\alpha}}\right)^\alpha + o\left(\frac{1}{n}\right)\right] = n\left[-\lambda\left(\frac{x}{n^{1/\alpha}}\right)^\alpha + o\left(\frac{1}{n}\right)\right] = -\lambda x^\alpha + o(1)$$

leading to $\lim_{n\to\infty} P\left(n^{1/\alpha} X_{(1)} > x\right) = e^{-\lambda x^\alpha}$, which completes the demonstration.

In fact, this property is at the origin of the form of the distribution suggested by W. Weibull in its 1951 article. He indeed proposes to solve problems of materials stiffness. The example which he uses to illustrate its distribution is that of a chain. How can one establish the probability that a chain breaks? Its reasoning is that the chain will break if the weakest of its links breaks. Which

[9] Together with the laws of Fréchet and Weibull; cf. Planchet et *al*. [2011].

is similar to determining the distribution of the minimum of a large number of objects. In extreme value theory, one establishes that the distribution of the minimum does not depend on the probability density function of each object, provided the number of objects is large enough (Galambos [1978], Gumbel [1958]). Logan [1992] used this distribution in the context of a race (race model). Let us imagine for example a large number of neurons in competition to emit a signal. The emitted signal will be produced by the fastest neuron.

One can finally observe that, since $S(t) = e^{-\lambda t^{\alpha}}$, one has $\mathbf{ln}(-\mathbf{ln}(S(t))) = \mathbf{ln}(\lambda) + \alpha\,\mathbf{ln}(t)$; if $\hat{S}(t)$, denotes the empirical survival function, the points $\left(\mathbf{ln}(t), \mathbf{ln}\left(-\mathbf{ln}\left(\hat{S}(t)\right)\right)\right)$ must thus be roughly aligned. This provides a simple way of checking whether duration data can be modelled by a Weibull distribution.

A different parameterisation of the Weibull distribution is sometimes used with $S(x) = \mathbf{exp}\left\{-\left(\frac{x}{l}\right)^{\alpha}\right\}$, which amounts to making the parameter shift $\lambda = l^{-\alpha}$. That also amounts to making the variable shift $y = \frac{x}{l}$ therefore modifying the unit of time used.

The model parameters' estimate is made by observing that likelihood is written:

$$L(\alpha,l) \propto \left(\frac{\alpha}{l^{\alpha}}\right)^{n} \prod_{i=1}^{n} t_i^{(\alpha-1)} \mathbf{exp}\left\{-\left(\frac{t_i}{l}\right)^{\alpha}\right\}$$

$$L(\alpha,l) \propto \left(\frac{\alpha}{l^{\alpha}}\right)^{n} \mathbf{exp}\left\{-l^{-\alpha}\sum_{i=1}^{n} t_i^{\alpha}\right\} \mathbf{exp}\left\{(\alpha-1)\sum_{i=1}^{n}\mathbf{ln}\, t_i\right\}$$

The following expression of log-likelihood is deduced:

$$\mathbf{ln}\, L(\alpha,l) = \mathbf{ln}\, k + n\left(\mathbf{ln}\,\alpha - \alpha\,\mathbf{ln}\, l\right) - l^{-\alpha}\sum_{i=1}^{n} t_i^{\alpha} + (\alpha-1)\sum_{i=1}^{n}\mathbf{ln}\, t_i$$

Partial differential equations are thus written:

$$\begin{cases} \dfrac{\partial}{\partial l}\mathbf{ln}\, L(\alpha,l) = -\dfrac{n\alpha}{l} + \alpha l^{-\alpha-1}\displaystyle\sum_{i=1}^{n} t_i^{\alpha} \\ \dfrac{\partial}{\partial \alpha}\mathbf{ln}\, L(\alpha,l) = n\left(\dfrac{1}{\alpha} - \mathbf{ln}\, l\right) + l^{-\alpha}\left(\mathbf{ln}\, l\displaystyle\sum_{i=1}^{n} t_i^{\alpha} - \sum_{i=1}^{n} t_i^{\alpha}\,\mathbf{ln}\, t_i\right) + \sum_{i=1}^{n}\mathbf{ln}\, t_i \end{cases}$$

One thus seeks the solutions of the following system:

$$\begin{cases} l = \left(\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} t_i^{\alpha}\right)^{1/\alpha} \\ \dfrac{1}{\alpha} = \dfrac{\displaystyle\sum_{i=1}^{n} t_i^{\alpha}\,\mathbf{ln}\, t_i}{\displaystyle\sum_{i=1}^{n} t_i^{\alpha}} - \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\mathbf{ln}\, t_i \end{cases}$$

The second equation can be solved numerically using a Newton-Raphson type algorithm which converges towards $\hat{\alpha}$, provided the initial seed is not too distant. Thus while noting $\varphi(\alpha) = \frac{\sum_{i=1}^{n} t_i^{\alpha} \ln t_i}{\sum_{i=1}^{n} t_i^{\alpha}} - \frac{1}{n}\sum_{i=1}^{n} \ln t_i - \frac{1}{\alpha}$, one will use the following recurrence relation:

$$\alpha_{i+1} = \alpha_i - \frac{\varphi(\alpha_i)}{\varphi'(\alpha_i)}$$

In practice, this value could be the estimator obtained by the quantiles method on the set of the complete observations, by observing that:

$$\breve{\alpha} = \frac{\ln\left[\frac{-\ln(1-p_2)}{-\ln(1-p_1)}\right]}{\ln\left(Q_{p_2}\right) - \ln\left(Q_{p_1}\right)}$$

with $Q_p = F^{-1}(p)$ the quantile function at point *p*. Reminder: any cumulative distribution function admits a generalised inverse function defined by:

$$F^{-1}(p) = \inf\left\{x\,;F(x) \geq p\right\}.$$

In the case of the Weibull distribution, one gets:

$$F^{-1}(p) = l\left[-\ln(1-p)\right]^{1/\alpha}$$

Once $\hat{\alpha}$ is obtained, $\hat{l}$ is deduced from the first equation.

3.3. Gamma model

The Gamma model is another natural generalisation of the exponential model: let us suppose that the duration T_r is the waiting time for the realisation of a service in a queue and that the queue is made up of *r* waiters, independent and identical, which each process one part of the service (they are therefore assembled in series). The assumption is made that the realisation duration of each waiter's process is an exponential law of parameter $\lambda > 0$.

Then the total duration of service is the sum of *r* exponential variables of the same parameter; the duration of service is therefore distributed according to a Gamma distribution of parameter (r,λ):

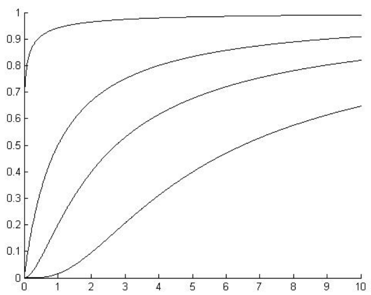
$$S_r(t) = \int_t^{+\infty} \frac{\lambda^r u^{r-1}}{(r-1)!} e^{-\lambda u} du.$$

<u>Demonstration</u>: if $L(a) = E(e^{-aT})$ indicates the Laplace transform at point *a* of an exponential law, then $L(a) = \frac{\lambda}{\lambda + a}$ and thus the Laplace transform of T_r is $L_{T_r}(a) = \left(\frac{\lambda}{\lambda + a}\right)^r$; which is a Laplace transform at point *a* of a Gamma distribution[10].

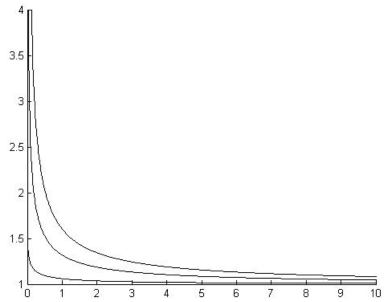
This law is called, when *r* is an integer, the Erlang distribution; one can define in the same way a duration model with a Gamma distribution which parameter *r* is not an interger[11]. The hazard function is as follows:

$$h(t) = \frac{t^{r-1}e^{-\lambda t}}{\int_t^{+\infty} x^{r-1}e^{-\lambda x}dx}.$$

The shape of this function is determined by the position of *r* compared to 1:



$r > 1$



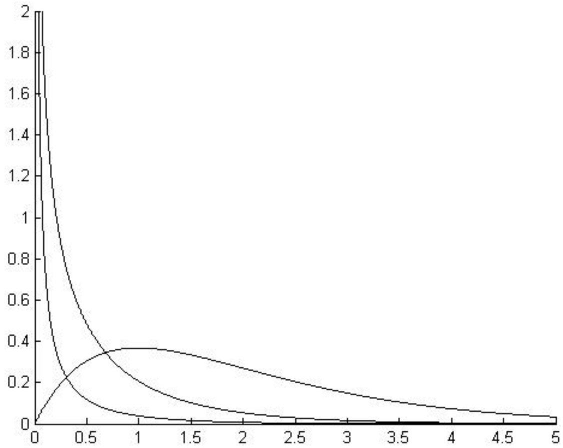
$r < 1$

The graphs above highlight in particular the fact that the Gamma distribution is not adapted *a priori* to the modelling of human mortality. Very fast decrease of the exit rate when $r < 1$ can on the other hand prove in line with the behaviour of duration in sick leave. The shape of this distribution is determined by the value of *r*; according to various values of *r* one obtains the graph below[12]:

[10] This is verified by a variable shift in the integral.

[11] With $r = \frac{n}{2}$ and $\lambda = 1/2$ one obtains the law of Khi-deux with *N* degrees of freedom.

[12] The graph represents the density of the Gamma law.



The mean and the variance of a Gamma distribution are given by:

$$E(T) = \frac{r}{\lambda} \text{ et } V(T) = \frac{r}{\lambda^2}.$$

One deduces from these expressions that the coefficient of variation of a Gamma distribution is:

$$cv = \frac{\sigma(T)}{E(T)} = \frac{1}{\sqrt{r}}.$$

One can thus very simply obtain a rough estimate of the parameter of form *r* by computing the reverse of the square of the coefficient of variation.

One can also check that the hazard function $h_{r,\lambda}$ is increasing if $r > 1$ and decreasing if $r < 1$; moreover $\lim_{t\to+\infty} h_{r,\lambda}(t) = \lambda$, which means that asymptotically we have the exponential model again.

Demonstration: by carrying out the variable shift *u*=*x*-*t* in the expression of the inverse hazard function, it becomes:

$$\frac{1}{h(t)} = \int_0^{+\infty} g(t,u) e^{-\lambda u} du$$

with $g(t,u) = \left(1 + \frac{u}{t}\right)^{r-1}$. The outcome results from the study of the sign of $\frac{\partial g}{\partial t}(t,u)$.

3.4. Gompertz-Makeham model

This is the model of reference for the construction of mortality tables. It is defined by the following hazard function:

$$h(t) = \alpha + \beta \times \gamma^t.$$

In demography, the form of this function is interpreted in the following way: the parameter α represents a rate of accidental death (independent of the age), while the term $\beta \times \gamma^t$ models an exponential ageing (if $\gamma > 1$). Incidentally we fall back on the exponential model if $\beta = 0$.

Compared to other models, the Makeham function therefore has an “explanatory” ambition, or “physical”, by explicitly integrating two clearly identified causes of death.

More precisely, if it is considered that death can occur from two “concurrent” causes, say accident and age, the timing of death is $T = T_A \wedge T_V$, T_A (resp. T_V) representing the accidental death (resp. due to age). One assumes that the accidental death is modelled by an exponential law of parameter α, and the death associated with ageing by the Gompertz hazard function $h(t) = \beta \times \gamma^t$; then T follows a Makeham distribution. This results from the fact that the survival function of T is the product of the survival functions of T_A and T_V, and thus the hazard functions are added.

A direct calculation easily leads to the expression of the survival function:

$$S_\theta(t) = \mathbf{exp}\left\{-\alpha t - \frac{\beta}{\mathbf{ln}(\gamma)}\left(\gamma^t - 1\right)\right\}.$$

However, the calculation of the expectancy of T is complex: $E(T) = \int_0^{+\infty} e^{-\alpha t - \frac{\beta}{\mathbf{ln}(\gamma)}\left(\gamma^t - 1\right)} dt$. But $S(t) = e^{\frac{\beta}{\mathbf{ln}(\gamma)}} \times e^{-\alpha t} \times e^{-\frac{\beta}{\mathbf{ln}(\gamma)}\gamma^t}$; the variable shift is then carried out:

$$u = \frac{\beta}{\mathbf{ln}(\gamma)}\gamma^t = \frac{\beta}{\mathbf{ln}(\gamma)} e^{t \times \mathbf{ln}(\gamma)}, \quad \frac{du}{\mathbf{ln}(\gamma) u} = dt$$

which implies $\left(\frac{\mathbf{ln}(\gamma)}{\beta} u\right)^{1/\mathbf{ln}(\gamma)} = e^t$ then:

$$E(T_\theta) = e^{\frac{\beta}{\mathbf{ln}(\gamma)}} \times \int_{\frac{\beta}{\mathbf{ln}(\gamma)}}^{+\infty} \left(\frac{\mathbf{ln}(\gamma)}{\beta} u\right)^{-\alpha/\mathbf{ln}(\gamma)} \times e^{-u} \frac{du}{\mathbf{ln}(\gamma) \times u}$$

$$= \frac{1}{\mathbf{ln}(\gamma)} \left(\frac{\mathbf{ln}(\gamma)}{\beta}\right)^{-\alpha/\mathbf{ln}(\gamma)} e^{\frac{\beta}{\mathbf{ln}(\gamma)}} \times \int_{\frac{\beta}{\mathbf{ln}(\gamma)}}^{+\infty} u^{-(1+\alpha/\mathbf{ln}(\gamma))} e^{-u} du$$

With the change of variable $v = u - \frac{\beta}{\mathbf{ln}(\gamma)}$ one finds:

$$E(T) = \frac{1}{\beta} \times \int_0^{+\infty} \left(\frac{\mathbf{ln}(\gamma)}{\beta} v + 1\right)^{-(1+\alpha/\mathbf{ln}(\gamma))} \times e^{-v} dv$$

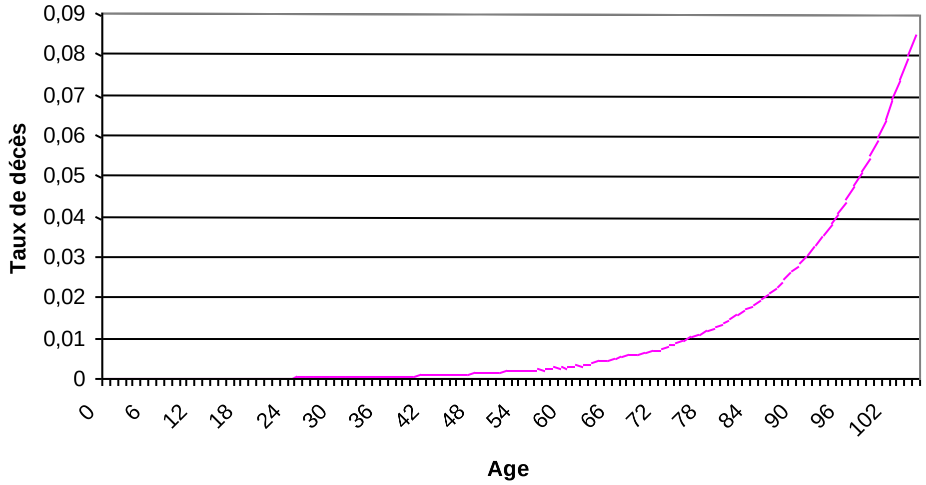
The expression above is complex and one can use the following simplified expression:

$$E_\theta(T) \approx e(\alpha, \beta, \gamma) = \sum_{t>0} \mathbf{exp}\left\{-\alpha t - \frac{\beta}{\mathbf{ln}(\gamma)}\left(\gamma^t - 1\right)\right\}.$$

Using the following "standard" parameters values for human mortality:

α	β	γ
8,81E-06	3,83E-05	1.076207

One finds the following shape of hazard rates depending on age:



One can graphically note the growth of the instantaneous rate of death with age – faster than in the case of a Weibull distribution (see 3.2 above) – which is in general better suited to human mortality.

One can finally observe that this model presents a geometrical property allowing, similarly to the case of a Weibull model, to graphically validate its adequacy with the data. Indeed, with $s = \exp(-\alpha)$ and $g = \exp\left(-\frac{\beta}{\ln(\gamma)}\right)$ as well as by observing that $-q_x \approx \ln(1-q_x) = \ln(s) + \gamma^x (\gamma - 1)\ln(g)$, one gets:

$$\ln(q_{x+1} - q_x) \approx x\ln(\gamma) + \ln\left((\gamma-1)^2 \ln(g)\right).$$

Under the assumption that mortality rates follow a Makeham distribution, the points $(x, y = \ln(q_{x+1} - q_x))$ are thus aligned on a line of slope $\ln(\gamma)$. The practical use of this remark will be developed further on.

4. Composite models

The object of this section is to describe the main features of basic models generally used within a parametric or semi-parametric framework, and using a sophistication degree higher than the simple analysis of an iid sample of parametric distribution fixed *a priori*. These are models generally used when dealing with a heterogeneous population, composed of individuals with various laws of survival; one has thus chosen to refer to these models as "composite models", and they differ by the way in which heterogeneity is taken into account.

The purely nonparametric models are studied in another lecture; they are not brought up here.

4.1. Mixture distributions

4.1.1. Introductory example

One considers a system made up of two independent elements assembled in parallel, each element's lifetime following an exponential distribution, with parameters λ_1 and λ_2.

The lifetime of the system is measured by $T = T_1 \vee T_2$; the distribution of T is obtained easily by observing that $1 - S(t) = \left(1 - e^{-\lambda_1 t}\right)\left(1 - e^{-\lambda_2 t}\right)$. One deduces that in the general case the hazard function is initially increasing, then decreasing; if $\lambda_1 = \lambda_2$, the hazard function is increasing. Temporal independence is thus a non-stable property and it is quickly lost.

One will see that it is also lost in the case of the aggregation of distributions.

4.1.2. Aggregation of distributions

Practically, it often happens that the observed durations result from the aggregation of subpopulations with each a specific behaviour, often unobservable. One then speaks of heterogeneity.

Let us assume that the survival function depends on a random parameter ν, this parameter being distributed according to π. From a heuristic point of view, one is in the presence of subpopulations within which the survival distribution is homogeneous and is described by the survival distribution conditional to the fact that the value of the parameter is ν, $S(t,\nu)$, π describing the respective weight of each subpopulation in the total population.

The initial survival function of the total population is therefore: $S(t) = \int S(t,\nu)\pi(d\nu)$:

$$S(t) = P(T > t) = E_\nu\left[P(T > t|\nu)\right] = \int S(t,\nu)\pi(d\nu).$$

The distribution of heterogeneity depends *a priori* on *t*, since the individuals of the various subpopulations do not leave the group at the same pace. At time *t*, and assuming an infinite population size, one gets:

$$\pi_t(d\nu) = \frac{S(t,\nu)}{S(t)}\pi(d\nu).$$

The hazard function at time *t* is then written $h(t) = \int h(t,\nu)\pi_t(d\nu)$. Indeed, one has to notice that:

$$u^{-1}P(T \le t+u|T > t) = \int u^{-1}P(T \le t+u|T > t,\nu)\pi_t(d\nu),$$

then to push *u* towards 0. In the particular case where $S(t,\nu) = \mathbf{exp}(-\lambda(\nu)t)$, i.e. where each subpopulation is described by an exponential law of parameter $h(t,\nu) = \lambda(\nu)$, the aggregate survival function is written:

$$S(t) = \int_0^{+\infty} \mathbf{exp}(-\lambda(\nu)t)\pi(d\nu)$$

According to the above expression, the hazard function is thus written $h(t) = \int \lambda(v)\pi_t(dv)$ and we deduce that:

$$\frac{dh(t)}{dt} = -\int \lambda^2(v)\pi_t(dv) + \left(\int \lambda(v)\pi_t(dv)\right)^2 .$$

Indeed, from $\pi_t(dv) = \frac{S(t,v)}{S(t)}\pi(dv)$ yields:

$$\frac{\partial}{\partial t}\pi_t(dv) = \frac{\frac{\partial}{\partial t}S(t,v) \times S(t) - S(t,v) \times \frac{d}{dt}S(t)}{S(t)^2}\pi(dv)$$

with $\frac{\partial}{\partial t}S(t,v) = -\lambda(v)S(t,v)$ and $\frac{S'(t)}{S(t)} = -h(t) = -\int \lambda(v)\pi_t(dv)$. We deduce that:

$$\frac{\partial}{\partial t}\pi_t(dv) = \frac{-\lambda(v) \times S(t,v)}{S(t)}\pi(dv) + \frac{S(t,v) \times h(t)}{S(t)}\pi(dv) = -\lambda(v)\pi_t(dv) + h(t)\pi_t(dv)$$

And writing $\frac{d}{dt}h(t) = \int \lambda(v)\frac{\partial}{\partial t}\pi_t(dv)$ we finally find:

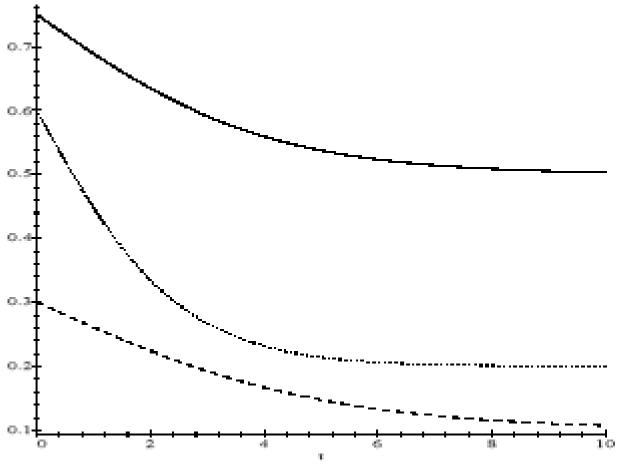
$$\frac{dh(t)}{dt} = -\int \lambda^2(v)\pi_t(dv) + h(t)^2 .$$

Which is the expected result. This equality implies by the Schwarz inequality (or by noticing that $\frac{dh(t)}{dt} = -V_{\pi_t}(\lambda(v))$) that $\frac{dh(t)}{dt} \leq 0$; the aggregation of constant hazard functions thus leads to a decreasing aggregated hazard function. This phenomenon is explained by the fact why the individuals having a high $\lambda(v)$ value exit first, leaving proportionally more individuals with low $\lambda(v)$ values with time. The exit rate logically decreases.

This phenomenon is called “heterogeneity bias”, or “mobile-stable”.

Example: mixture of two exponential laws

Here, duration is an exponential variable of parameter λ_1 with probability p and λ_2 with probability 1-p, that is to say $S(t) = pe^{-\lambda_1 t} + (1-p)e^{-\lambda_2 t}$. The hazard function takes the following form:



It shows that the instantaneous risk can quickly decrease, while the two origin functions remain at constant risk.

4.2. Proportional hazards models

These are semi-parametric models with a basis survival function $B(t)$, and the assumption that the survival function of the observed phenomenon is of the form $S_\theta(t) = B(t)^\theta$, for an unknown parameter $\theta > 0$. The underlying density is written $f_\theta(t) = \theta B(t)^{\theta-1} f(t)$ and the hazard function is thus of the form:

$$h_\theta(t) = \frac{f_\theta(t)}{S_\theta(t)} = \theta \frac{f(t)}{B(t)} = \theta h(t)$$

The hazard function is therefore proportional to the basis hazard function associated with $\theta = 1$, thus the name "proportional hazard model". The exponential model is a particular case of proportional hazard model in which the basis hazard function is constant and equal to one.

One can notice that these models satisfy the following property: if the random variable T_θ is associated with the survival function $S_\theta(t) = B(t)^\theta$, then $E(T_\theta) = \int_0^{+\infty} S_\theta(t)dt = \int_0^{+\infty} B(t)^\theta dt$; however one recognizes in $\rho_\theta(T) = \int_0^{+\infty} B(t)^\theta dt$ the Wang[13] risk measure associated with the distortion function $g_\theta(x) = x^\theta$ (called *PH-transform* of parameter $\frac{1}{\theta}$).

By specifying different forms for the proportionality factor, one is led to define various classes of models.

[13] See for example Planchet *et al.* [2011] for a more general presentation of risk measures.

4.2.1. Cox model

This model can integrate determinants used to define the parameter $\theta > 0$; for this one writes $\theta = e^{z'\beta}$ with $z = (z_1, \ldots, z_p)$ a vector of p determinants and $\beta = (\beta_1, \ldots, \beta_p)$ the vector of parameters; with this formulation one gets:

$$\mathbf{ln}\, h(t|Z = z) = \mathbf{ln}\, h(t) + \sum_{i=1}^{p} z_i \beta_i$$

and thus a linear regression model. This model is called the Cox model. It can be apprehended in two different ways, according to whether the basis hazard function *h* is assumed known (for example assuming that it is a Weibull model) or unknown. In this latter case, it becomes a nuisance parameter of infinite size which makes the estimation of the other parameters complicated.

4.2.2. Fragility models

In Cox model one seeks to model the effect of known determinants on the level of the risk function; in some situations, these variables are unobservable, and still one wishes to assess their impact on the form of the survival function.

One starts again from $S_\theta(t) = S(t|\theta) = B(t)^\theta$ or, in an equivalent way, $h_\theta(t) = \theta h(t)$ of a proportional hazard model, and one considers that the parameter θ is a random variable; in other words the conditional (to the parameter) survival distribution is given, and the total distribution is thus obtained by integration:

$$S(t) = E\left[B(t)^\theta\right]$$

the expectancy being calculated according to the distribution of θ. This expression is similar to the expression $S(t) = \int S(t, \nu)\pi(d\nu)$ obtained in section 4.1.2. The parameter θ is called “fragility”. These models are also sometimes called “random effects models”.

Traditional approach

The fragility models were introduced by Vaupel et al. [1979] to give an account of individual heterogeneity in a mortality context. The parameter of fragility makes it possible in practice to introduce differences in level of mortality between individuals, by assuming that the evolution of mortality with age is identical for all individuals. Heterogeneity is then modelled *via* the distribution of the parameter θ. In Vaupel et al. [1979] the following assumption on the distribution $\gamma(r, \lambda)$ is made:

$$\pi(d\theta) = f_{r,\lambda}(\theta) = \frac{\lambda^r \theta^{r-1}}{\Gamma(r)} \mathbf{exp}(-\lambda\theta)$$

that one chooses of mean 1, by imposing $r = \lambda$ and by considering as control parameter the variance $\sigma^2 = \lambda^{-1}$. In this case, and for a population observed since birth, one can show that the average hazard function of the population at the age *t* is of the form:

$$\bar{h}(t) = h(t) S(t)^{\sigma^2}$$

with $S(t)$ the number of survivors at age t. In this expression one has $\bar{h}(t) = \int h_\theta(t)\pi_t(d\theta) = h(t)\int \theta\pi_t(d\theta)$ with $\pi_t(d\theta) = \frac{S(t,\theta)}{S(t)}\pi(d\theta)$. In addition, $S(t) = \int S(t,\theta)\pi(d\theta)$.

This model was generalised by Barbi [1999] which proposed, by always assuming a proportional fragility initially distributed according to a Gamma distribution, a heterogeneity model called "combined fragility", in which besides the parameter θ, we have a discrete distribution τ independent of θ such as:

$$h_{\theta,\tau}(t) = \theta h(t,\tau).$$

This is equivalent to subdividing the initial population in sub-groups each described, conditionally to the proportional fragility factor θ, by an individual risk function. This model is noticeably used in Barbi and Al [2003] to study extreme survival ages. These authors write:

$$h(t,\tau_i) = a \times \mathbf{exp}(b_i \times x) + c$$

which is equivalent to assuming that global observed mortality is a mixture of Makeham distributions (still with Gamma assumption for the proportional fragility distribution). The aggregate risk function is then of the form:

$$\bar{h}(x) = \sum \pi_i(x) h(x,\tau_i) s_{x|\tau_i}(x,\tau_i)^{\sigma^2}$$

with $\pi_i(x)$ the proportion of individuals of group i surviving at age x.

Alternative approach

This modelling approach is also useful for introducing dependency between various lifetimes. For that purpose one assumes that the observed durations, $T_1,\ldots,T_n$ are independent conditionally to θ and that the marginal (conditional) distributions are of the form $S_i(t|\theta) = B_i(t)^\theta$; the joint survival function is directly deduced:

$$S(t_1,\ldots,t_n) = E\left[\left(B_1(t_1)\ldots B_n(t_n)\right)^\theta\right]$$

In this case the fragility parameter is interpreted as an exogenous element which modifies the behaviour of the whole group of individuals. In general, the basis survival function is identical for all the individuals and we have:

$$S(t_1,\ldots,t_n) = E\left[\prod_{i=1}^{n} B(t_i)^\theta\right]$$

But since $B(t) = \mathbf{exp}(-H(t))$ where H is the cumulated reference hazard function, this expression becomes:

$$S(t_1,\ldots,t_n) = E\left[\mathbf{exp}\left(-\theta\sum_{i=1}^{n} H(t_i)\right)\right]$$

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbb{1}_{]t;\infty[}(T_x)$$

One recognises on the right-hand side of the equation the Laplace transform of the variable θ at the point $\sum_{i=1}^{n} H(t_i)$. When θ is distributed according to a stable distribution of parameter α (i.e. the Laplace transform of θ is $E\,\mathbf{exp}(-x\theta) = \mathbf{exp}(-x^\alpha)$) one obtains the Hougaard model (*cf* Hougaard [2000]) with the survival function $S(t_1,\ldots,t_n) = \mathbf{exp}\left\{-\left[\sum_{i=1}^{n}\left(-\mathbf{ln}\,S(t_i)\right)^{1/\alpha}\right]^\alpha\right\}$; one can notice that the joint distribution being of the form $C(S_1,\ldots,S_n)$ a copula is thus defined, known as Hougaard copula.

4.3. Increasing transformations of duration

This is a semi-parametric model with a basis survival function, $S(t)$ and the assumption that the survival function of the observed phenomenon is $S_\theta(t) = S(\theta t)$, for a parameter $\theta > 0$. The hazard function is written:

$$h_\theta(t) = \frac{f_\theta(t)}{S_\theta(t)} = \theta \frac{f(\theta t)}{S(\theta t)} = \theta h(\theta t)$$

and this expression does not simplify like in the case of the proportional hazard model.

One can however notice that the two approaches are equivalent if and only if the hazard function is constant: indeed if the model is with proportional hazard one must find a hazard function k such that $h_\theta(t) = \theta k(t)$ and thus the functions k and h must satisfy the equality $k(t) = h(\theta t)$, which is possible only if the two functions are constant. Then we fall back within the framework of the exponential model.

This approach can be generalised provided we have an increasing function ψ_θ by considering the survival functions $S_\theta(t) = S\left(\psi_\theta^{-1}(t)\right)$; this is equivalent to studying the variables $\psi_\theta(T)$, where T is the basic variable. The Weibull distribution provides an example with $\psi_\alpha(t) = t^{1/\alpha}$ and an exponential law (see section 3.2).

4.4. Models with multiple exit causes

In some situations, one needs to differentiate various exit causes; for example, when studying mortality one is interested in the cause of death, *etc*. It is typically what is done when interpreting the Makeham model (see 3.4 above).

Let us note $T_1,..,T_n$ the variables associated with each studied cause, global survival is simply $T = T_1 \wedge .. \wedge T_n$; under the assumption of independence of the various components the model is simple and the total hazard function is the sum of hazard functions. But the assumption of independence can sometimes be restrictive, and the fragility models provide a simple way to get free from it. This approach was initially proposed by Oakes [1989].

It is thus assumed that the durations associated with each cause, $T_1,\ldots,T_n$ are independent conditionally to θ and that the marginal (conditional) distributions are of the form $S_i(t|\theta) = B_i(t)^{\theta}$. Follows a series of calculations similar to those in section 4.2.2 above and one finds:

$$S(t_1,\ldots,t_n) = E\left[\prod_{i=1}^{n} B_i(t_i)^{\theta}\right]$$

Example: with two causes of exit each distributed according to a Weibull distribution and a distribution of the mixture parameter according to a stable law of parameter *a*, one finds $S(t) = \mathbf{exp}\left\{-\left(\lambda_1 t^{\alpha_1} + \lambda_2 t^{\alpha_2}\right)^{a}\right\}$, who is an immediate consequence of $E(\mathbf{exp}(-x\theta)) = \mathbf{exp}(-x^{a})$ and of the expression of the survival function of the Weibull distribution $S(t) = \mathbf{exp}(-\lambda t^{\alpha})$.

4.5. Models with common shock

The idea here is that the survival duration depends on two factors, one related to the individual and the other affecting the population as a whole. This second factor can be an accidental or environmental factor. One considers the following model:

$$T_i = X_i \wedge Z$$

With S_i the survival function of X_i and S_Z the survival function of *Z*. The joint distribution of the vector $(T_1,\ldots,T_n)$ is obtained by observing the fact that the event $\{X_i \wedge Z > t\}$ is equal to $\{X_i > t\} \cap \{Z > t\}$, which leads to:

$$S(t_1,\ldots,t_n) = \prod_{i=1}^{n} S_i(t_i) \times S_Z(\mathbf{max}(t_1,\ldots,t_n)).$$

Marshall and Olkin [1967] propose for example an exponential distribution for *Z*.

5. References

BARBI E.[1999] « Eterogeneità della popolazione e sopravvivenza umana : prospettive metodologiche ed applicazioni alle generazioni italiane 1870-1895 », Florence, Dipartimento Statistico (Ph.D. thesis), 91 p.

BARBI E., CASELLI G. et VALLIN J. [2003] « Hétérogénéité des générations et âge extrême de la vie », *Population* 2003/1, Vol. 58, p. 45-68.

BÖHMER P.E. [1912], « Theorie der unabhängigen Wahrscheinlichkeiten », Rapports, Mém. et Procès-verbaux 7e Congrès Int. Act. Amsterdam 2, 327–343.

CAIRNS A., BLAKE D., DOWD K. [2004], « *Pricing Frameworks for Securitization of Mortality Risk* », AFIR

COX D.R. [1972] « Regression models and life-tables (with discussion) ». *J. R. Statist. Soc. Ser. B*, pages 187-220.DROESBEKE J.J., FICHET B., TASSI P. [1989], *Analyse statistique des durées de vie.* Paris : Economica

GALAMBOS J. [1978] *The Asymptotic Theory of Extreme Order Statistics*. New York : John Wiley and Sons.

GUMBEL E. J. [1958] *The Statistics of Extremes*. New York : Columbia University Press.

Hougaard P. [2000] *Analysis of Multivariate Survival Data*, New York : Springer-Verlag.

Kaplan E.L., Meier P. [1958] « Non-parametric estimation from incomplete observations ». *Journal of the American Statistical Association*, 53, 457-481.

Langmeier R. [2000], « Etude de différentes méthodes d'ajustement de tables de mortalité : application aux données d'une compagnie d'assurance », HEC Lausanne

Le Bras H. [2000] *Naissance de la mortalité*, Col. Hautes Etudes, Gallimard – Le Seuil

Lee, R.D., Carter, L. [1992] « Modelling and forecasting the time series of US mortality ». *Journal of the American Statistical Association* 87,659-671.Logan G. D. [1992] « Attention and preattention in theories of automaticity ». *American Journal of Psychology*, 105, 317-339.

Makeham W. [1860] « On the law of mortality. » Journal of the Institute of Actuaries, 13, 325-358.

Marshall AW., Olkin, I. [1967] « A generalized bivariate exponential distribution ». *J. Appl. Probability* 4, 291-302

Oakes D. [1989] « Bivariate survival models induced by frailties ». *Journal of the American Statistical Association* 84, 487-493.

Partrat C., Besson J.L. [2004], *Assurance non-vie - modélisation, simulation.* Paris : Economica

Planchet F., Thérond P.E. [2011] *Modélisation statistique des phénomènes de durée – applications actuarielles*, Paris : Economica.

Vaupel J. W., Manton K., Stallard E., 1979, « The impact of heterogeneity in individual frailty on the dynamics of mortality », *Demography*, 16, p. 439-454.

Weibull W. [1951] « A Statistical Distribution Function of Wide applicability ». *Journal of Applied Mechanic*, 18, p. 292-297.

6. Appendix: generally used Laplace transforms

Fonctions		Transformée de Laplace
dérivée	$\frac{d}{dt}f(t)$	$s\widehat{f}(s) - f(0)$
dérivée k^{ieme}	$\frac{d^k}{dt^k}f(t)$	$s^k\widehat{f}(s) - s^{k-1}f(0) - s^{k-2}f'(0) - \cdots - f^{(k-1)}(0)$
intégrale	$\int_0^t f(x)dx$	$\frac{1}{s}\widehat{f}(s)$
dilatation ($\lambda > 0$)	$f(\lambda t)$	$\frac{1}{\lambda}\widehat{f}(s/\lambda)$
translation	$f(t - t_0)$	$\exp(-st_0)\widehat{f}(s)$
facteur polynomial	$(-t)^k f(t)$	$\frac{d^k}{ds^k}\widehat{f}(s)$
facteur exponentiel	$\exp(-\lambda t)f(t)$	$\widehat{f}(s+\lambda)$
Heaviside	$Y(t)$	$1/s$
Dirac	$\delta(t)$	1
exponentielle	$\lambda e^{-\lambda t}Y(t)$	$\frac{\lambda}{s+\lambda}$
gamma	$\frac{\lambda^\beta t^{\beta-1}e^{-\lambda t}}{\Gamma(\beta)}Y(t)$	$\frac{\lambda^\beta}{(s+\lambda)^\beta}$
sinus	$\sin(\omega t)Y(t)$	$\frac{\omega}{s^2+\omega^2}$
cosinus	$\cos(\omega t)Y(t)$	$\frac{s}{s^2+\omega^2}$