STATISTICAL ANALYSIS OF FAILURES OF A REDUNDANT SYSTEM WITH ONE OPERATING UNIT AND ONE STAND-BY UNIT IN WARM OPERATING STATE

Vilijandas Bagdonavicius and Inga Masiulaityte
Vilnius University, Lithuania

Mikhail Nikulin
IMB UMR 5251, INSERM U593, University Victor Segalen, Bordeaux
France

Maine University, Le Mans , 23 March, 2007

PLAN

1. **Introduction and Notations**
2. **Accelerated Life or Flexible or Dynamic Regression Models**
3. **Explanatory variables, Covariates, Stresses**
4. **Sedyakin Principle and AFT Model**
5. **Statistical Analysis of Redundant System**

1. Introduction

We consider a **redundant system** with one operating unit and **one stand-by unit**, (additional structure, used for increasing the reliability of a system). **If** the main unit fails then the stand-by unit (if it is not failed yet) is commuted and operates instead of the main one. We suppose that **commuting is momentary** and there are no repairs. **If** the stand-by unit is not functioning until the failure of the main unit (**"cold" reserving**), it is possible that during and after commuting the **failure rate increases** because the stand-by unit is not "**warmed**" enough. **If** the stand-by unit is functioning in the same "**hot**" conditions as the main unit then usually after commuting the reliability of the stand-by unit does not change. **But** "**hot**" redundancy has **disadvantages** because the stand-by unit fails **earlier** than the main one with the probability 0.5.

So "**warm**" reserving is sometimes used: the stand by unit functions under **lower stress** than the main one.

In such a case the probability of the failure of the stand-by unit is **smaller** than that of the main unit and it is also possible that commuting is fluent. So the **main problem** is to **verify the hypothesis** that the switch on from "**warm**" to "**hot**" conditions does not do some **damage** to units.

The Data. Suppose that the following **data** are available :

a) the failure times $T_{11}, \dots, T_{1n_1}$ of n_1 units tested in **"hot" conditions**;

b) the failure times $T_{21}, \dots, T_{2n_2}$ of n_2 units tested **in "warm" conditions**;

c) the failure times $T_1, \dots, T_n$ of n **redundant systems** (with "**warm**" stand-by units).

We shall consider the model which gives the new statistical approach for determining the **basic reliability measures** of **redundant systems**. It is evident that the considered question is related with phenomena of **longevity, aging, fatigue** and **degradation** of complex systems. Statisticians working in survival analysis, reliability, biostatistics apply the so-called **dynamic regression models** which are well adapted to study such kinds of problems.

Remark on Aging and Longevity, Failure and Degradation

It is well known that **traditionally** only the **failure time data** are usually used for product reliability estimation or for estimation of survival characterictics. Failures of highly reliable units are **rare** and other information should be used in addition to censored failure-time data. **One way** of obtaining a complementary reliability information is to use **higher levels of experimental factors** or covariates (such as temperature, voltage or pressure) to **increase** the number of failures and, hence, to obtain reliability information quickly. The **accelerated life testing** of bio-technical systems is meant to be a simple practical method of estimation of the reliability of new systems without having to wait the operating life of them. It is evident that the **extrapolating reliability** from ALT always carries the risk that the **accelerating stresses does not** properly **excite** the **failure mechanism** which dominate at operating (**normal**) stresses.

Another way of obtaining this complementary reliability information is to measure some parameters (covariates) which characterize the **aging of the product in time**. In analysis of **longevity** of highly reliable complex industrial or biological systems, the degradation processes provide **additional information** about the **aging, degradation** and **deterioration** of systems, and from this point of view the **degradation data** are really a very rich source of reliability information and often offer many **advantages** over failure time data. **Degradation** is the **natural response** for some tests, and it is natural also that with degradation data it is possible to make useful reliability and statistical inference, even **with no failures**. It is evident that sometimes it may be difficult and costly to collect degradation measures from some components or materials. Sometimes it is possible to apply the **expert's estimation** of the level of degradation.

More about constructions and applications of **dynamic regression models** one can see, for example, in Yashin (2004), Nikulin, Commenges and Huber (2006), **Ceci and Mazliak (2004)**, Meeker and Escobar (1998), **Martinussen and Scheike** (**2006**), etc...
Reliability theory and **survival analysis** provide a great method to obtain a general theory of **aging and degradation of complex technical and bio-technical systems**, see, Bagdonavicius and Nikulin (2002), Limnios and Nikulin (2000), Nikulin, Balakrishnan, Limnios and Mesbah (2004), Scheike (2006), Lindqvist and Doksum (2003), Huber (2006), Huber, Vonta and Solev (2006), **Zeng** and **Lin** (2007), Vonta, Nikulin, Huber, Limnios (2007), Huber, Limnios, Mesbah and Nikulin (2007), Wu (2004), Dabrowska (2005,2006), etc.

Notations. Suppose that T is the random **time-to-failure** of an **unit** (or system). We say also that T is a **hard** or **traumatic failure**. Let $S(\cdot)$ be the **survival function** and $\lambda(\cdot)$ be the **hazard rate**:

$$S(t) = \mathbf{P}\{T > t\}, \quad \lambda(t) = \lim_{h \to 0} \frac{1}{h} \mathbf{P}\{t \le T < t+h | T \ge t\} = -\frac{d[\ln S(t)]}{dt},$$

from where it follows that $S(\cdot)$ can be written as

$$S(t) = e^{-\Lambda(t)}, \quad \text{where} \quad \Lambda(t) = \int_0^t \lambda(s) ds$$

is the **cumulative hazard function**. $F(\cdot) = 1 - S(\cdot)$ is the **cumulative distribution function** of T. In **Survival Analysis** and **Reliability** the models formulated in terms of cumulative hazard and hazard rate functions. The **most common shapes** of hazard rates are **monotone**, $\cup$-**shaped** or $\cap$-**shaped** (Viertl (1988), Meeker and Escobar (1998), Bagdonavicius and Nikulin (2002)).

Note that, when $\lambda(t) \equiv \lambda = const > 0$, the system is **nonaging**. This case corresponds to the **parametric exponential distribution** of the T:

$$S(t) = \mathbf{P}\{T > t\} = e^{-\lambda t}, \quad t > 0, \quad (\lambda > 0),$$

with

$$\mathbf{E}T = \frac{1}{\lambda}, \quad \mathbf{Var}T = \frac{1}{\lambda^2}$$

and the **linear cumulative hazard function**

$$\Lambda(t) = \lambda t, \quad t > 0.$$

From the condition $\lambda(t) \equiv \lambda$ it follows that for any $t, s > 0$

$$\mathbf{P}\{T > t + s | T > s\} = \mathbf{P}\{T > t\} = e^{-\lambda t}.$$

We say that there is the property of the **lack-of memory**.

The more classical distributions for hazard rate are the next.

1. **Gompertz-Makeham law**, with

$$\lambda(t) = \beta + \alpha e^{\gamma t}, \quad t > 0, \alpha > 0, \gamma > 0, \beta > 0.$$

2. The **Weibull law**, with

$$S(t) = e^{-(t/\theta)^{\nu}}, \quad \lambda(t) = \frac{\nu}{\theta^{\nu}} t^{\nu-1}, \quad t > 0, \nu > 0, \theta > 0.$$

3. The **Gamma-Model** with

$$\lambda(t) = \frac{t^{p-1} e^{-\lambda t}}{\int_t^{\infty} u^{p-1} e^{-\lambda u} du}, \quad t > 0, p > 0, \lambda > 0.$$

More about **stochastic parametric models** one can see Nelson (1990), Meeker and Escobar (1998), Limnios and Nikulin (2000), Lawless (2003), Bagdonavicius and Nikulin (1995), Yashin (2004), Duchesne (2004), Lehmann (2004,**2006**), Kahle (**2004, 2006**), Nikulin, Balakrishnan, Limnios, Mesbah (2004), etc...

2. Accelerated life or Flexible or Dynamic Regression Models

The main objective of the **accelerated life testing** (**ALT**) is to obtain **more information** concerning failures from a given test time under **accelerating** stresses **than** would **normally be possible**. In industrial sector this information should then be used in order to quantify reliability and improve product reliability efficiently, and to determine whether safety and reliability goals are being meet, to assess warranty risk, etc...(see, Meeker and Escobar (1998), Lehmann (2004)).

Accelerated life models relating the lifetime distribution to possibly **time dependent explanatory variables** are considered here, see, for example, Dabrowska (2005), Lawless (2003), Duchesne (2004), Lin and Ying (1995), Singpurwalla (1971), Viertl (2004), Nelson (1990), Kahle (2006), Lehmann (2006), Wu (2006), etc...

3. Explanatory variables, Covariates, Stresses

Suppose that the **explanatory variable** is a **deterministic** time function

$$x(\cdot) = (x_1(\cdot), \ldots, x_m(\cdot))^T : [0, \infty[\rightarrow B \in R^m,$$

which is a vector of **covariates** itself or a **realization** of a **stochastic process** $X(\cdot) = (X_1(\cdot), \ldots, X_m(\cdot))^T$, which is called also the **covariate process** . We denote E a **set** of all **possible** (**admissible**) **stresses**. If $x(\cdot) \in E$ is **constant in time** we denote x instead, and we note E_1 the **set** of all constant covariates, $E_1 \subset E$.

The covariates can be interpreted as the **control** (see Ceci and Mazliak, (2004)), since we may consider **models of aging** in terms of differential equation and so to use all theory and techniques from the **optimal control theory**. We may say that we consider statistical **modelling with dynamic design** or **in dynamic environments**.

Step-Stresses

The mostly used **time-varying stresses** in **ALT** are **step-stresses**: units are placed on test at an initial low stress and if they do **not fail** in a predetermined time t_1, the stress is **increased**. If they do not fail in a predetermined time $t_2 > t_1$, the stress is **increased** once more, and so on. Thus **step-stresses** with k steps have the form

$$x(t) = x_1 \mathbf{1}_{\{0 \le t < t_1\}} + x_2 \mathbf{1}_{\{t_1 \le t < t_2\}} + ... + x_k \mathbf{1}_{\{t_{k-1} \le t < t_k\}}, \qquad (1)$$

where $x_1, \cdots, x_k$ are from E_1, $\quad t_k \le \infty$. **Sets** of all **possible step-stresses** of the form (1) will be denoted by E_k, $E_k \subset E$.

Let E_2, $E_2 \subset E$, be a **set** of step-stresses of the form

$$x(t) = x_1 \mathbf{1}_{\{0 \le t < t_1\}} + x_2 \mathbf{1}_{\{t_1 \le t\}}, \quad x_1, x_2 \in E_1. \tag{2}$$

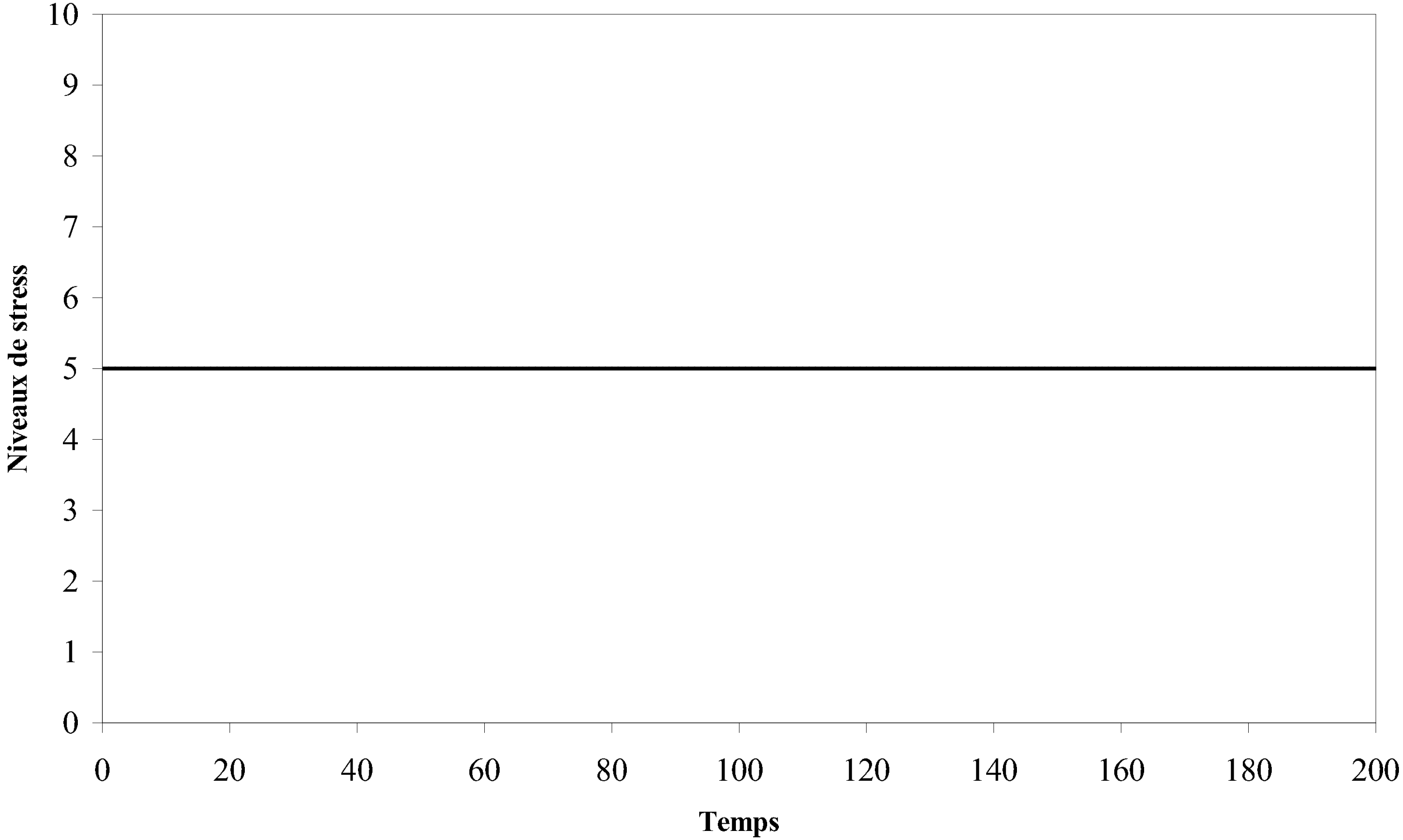
Let $T_{x(\cdot)}$ be the **failure time under** a stress $x(\cdot) \in E$. It is evident that if our **data are censored**, then we have to consider the influence of $x(\cdot)$ on the distribution of censoring time C, i.e. we have to write $C = C_{x(\cdot)}$, and hence our **observation** is

$$X_{x(\cdot)} = \min(T_{x(\cdot)}, C_{x(\cdot)}).$$

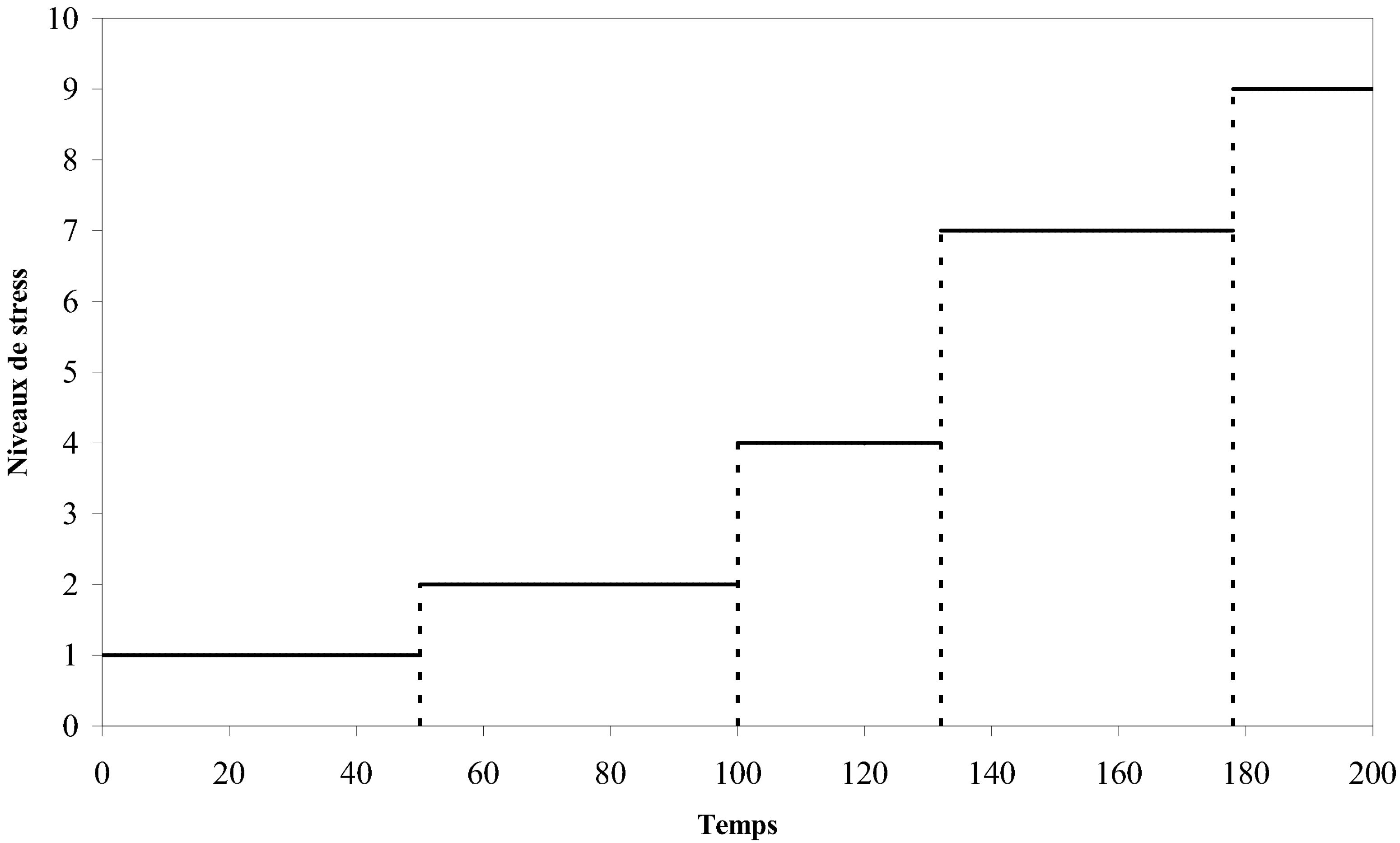
Before to speak about the construction of the models

$$\{S_x, x \in E_1\}, \{S_{x(\cdot)}, x(\cdot) \in E_2\}, \{S_{x(\cdot)}, x(\cdot) \in E_k\}, \{S_{x(\cdot)}, x(\cdot) \in E\}$$

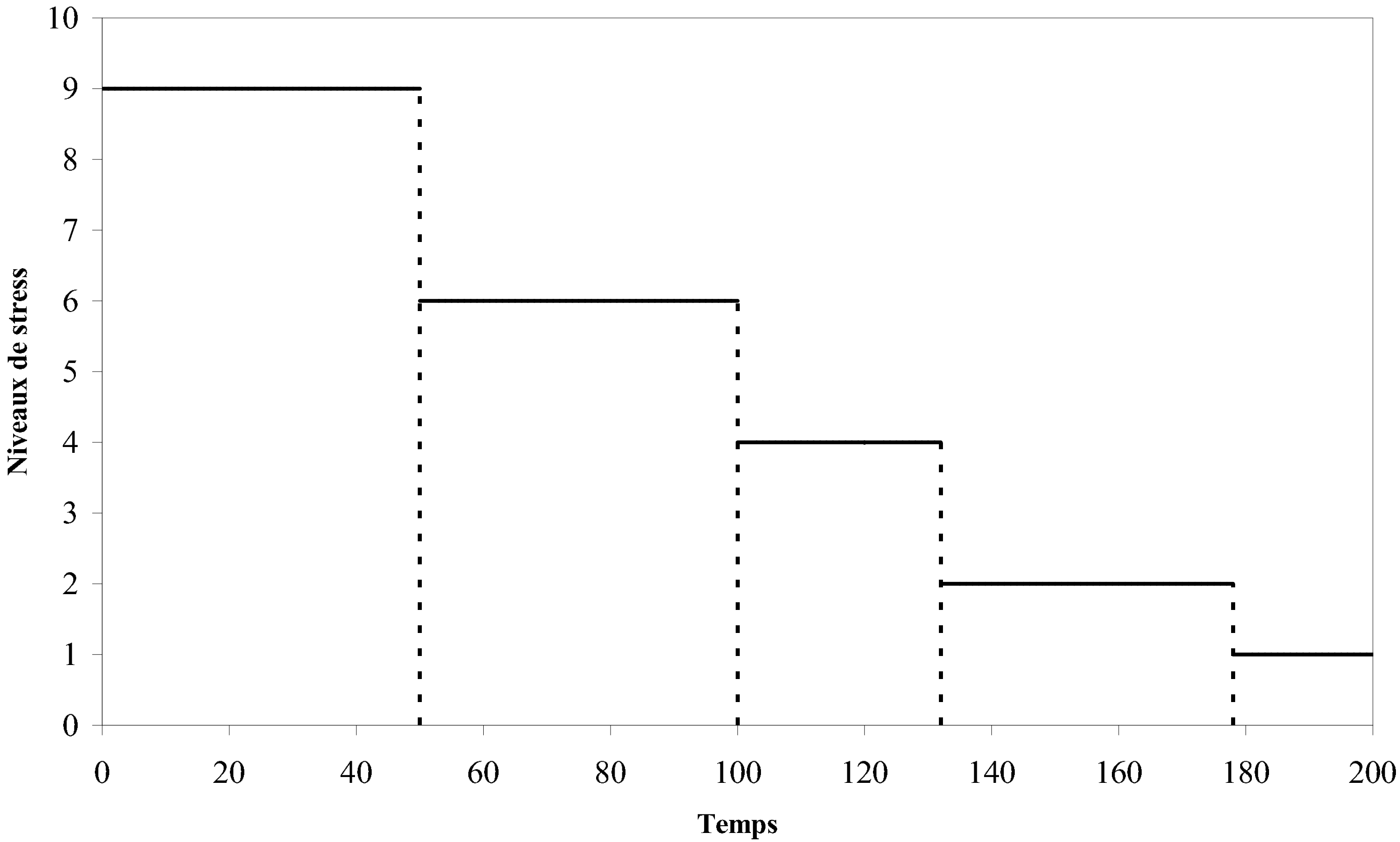
on E_1, E_2,...,E_k,, or on E, where $S_{x(\cdot)}$ is the **survival** function of T given $x(\cdot)$, we shall give here several pictures of possible **scalar** stresses ($m = 1$).



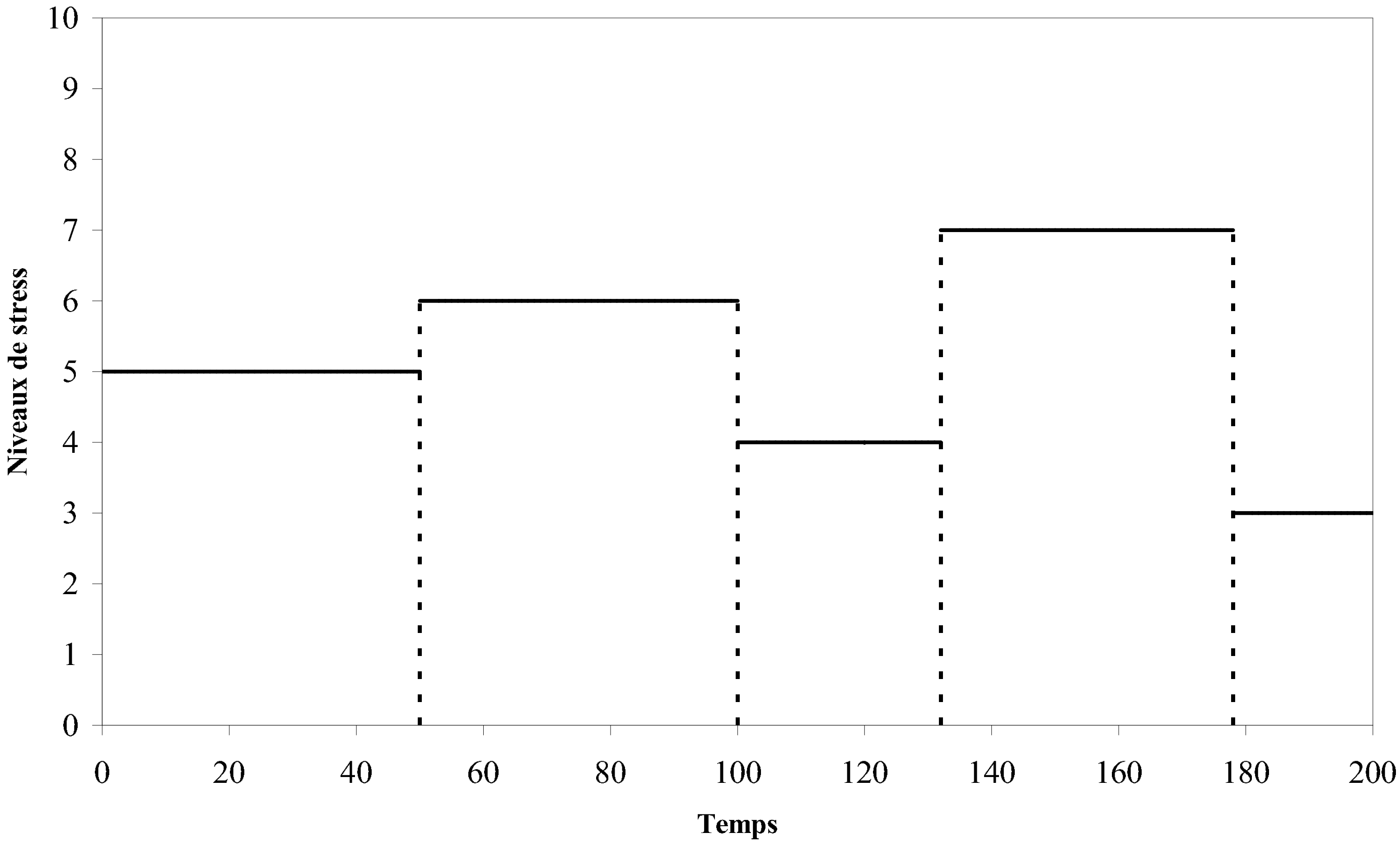
Stress $x = 5$ is a constant in time, $x \in E_1$.



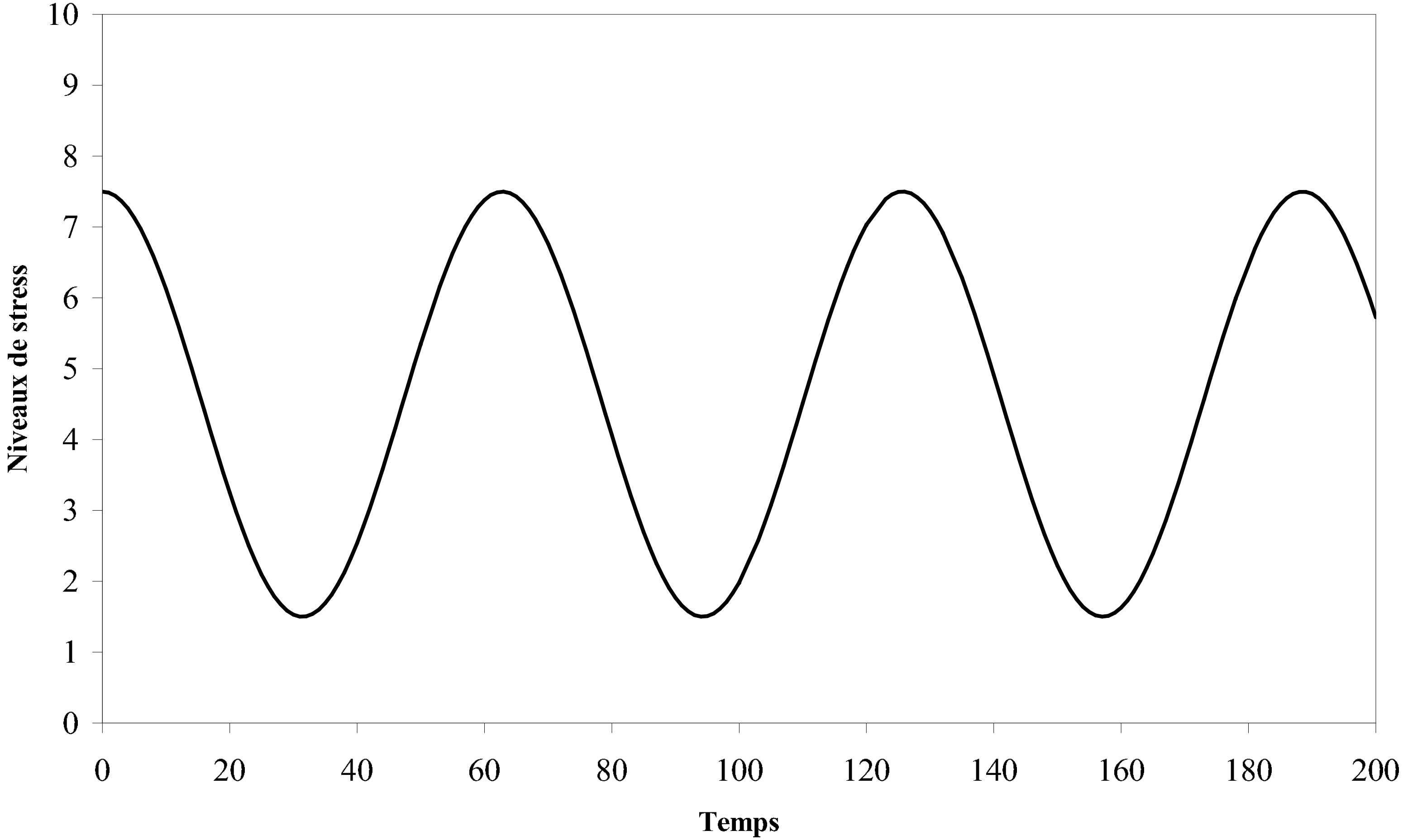
Stress x is a increasing step stress, $x \in E_5$.



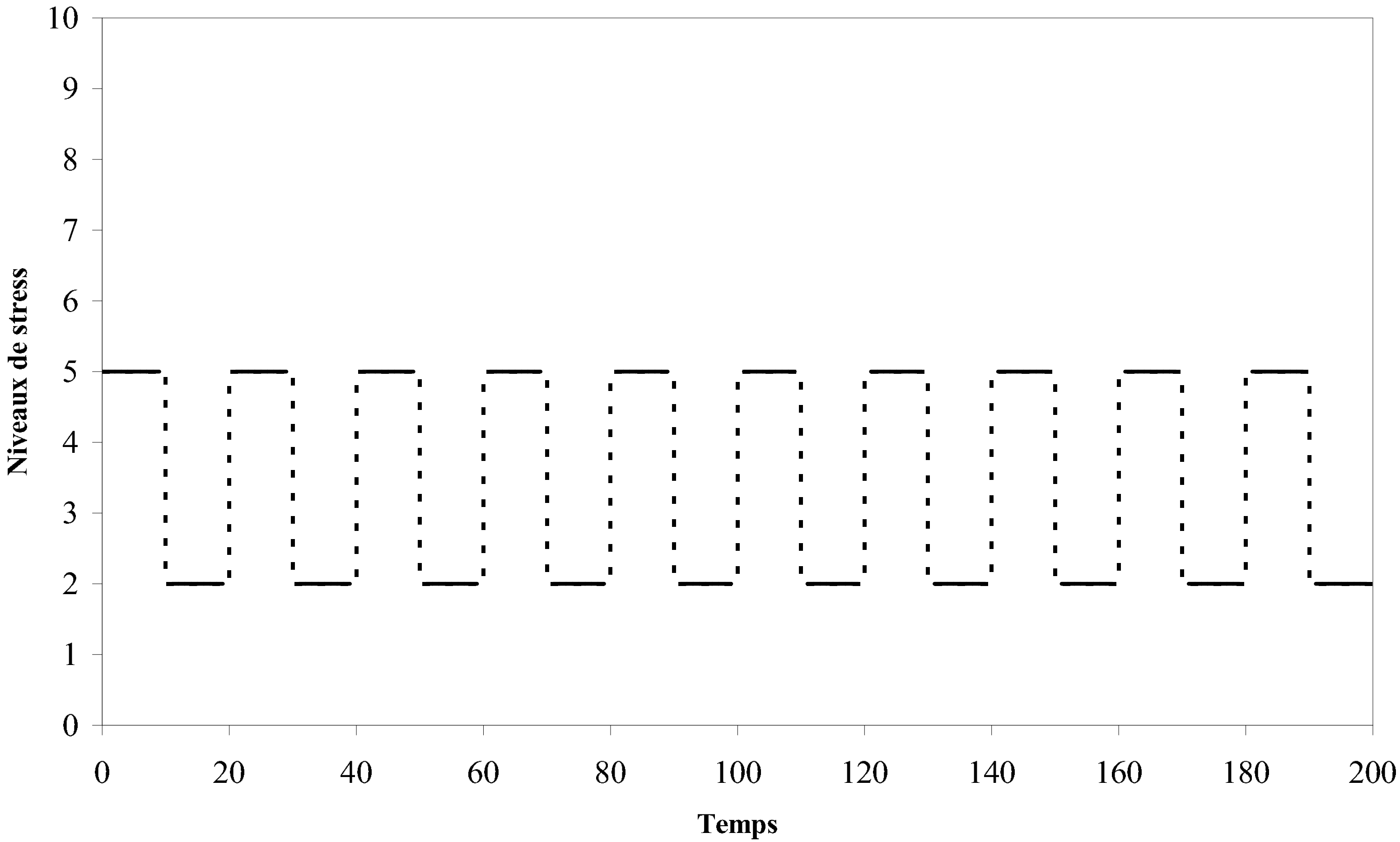
Stress x is a decreasing step stress, $x \in E_5$.



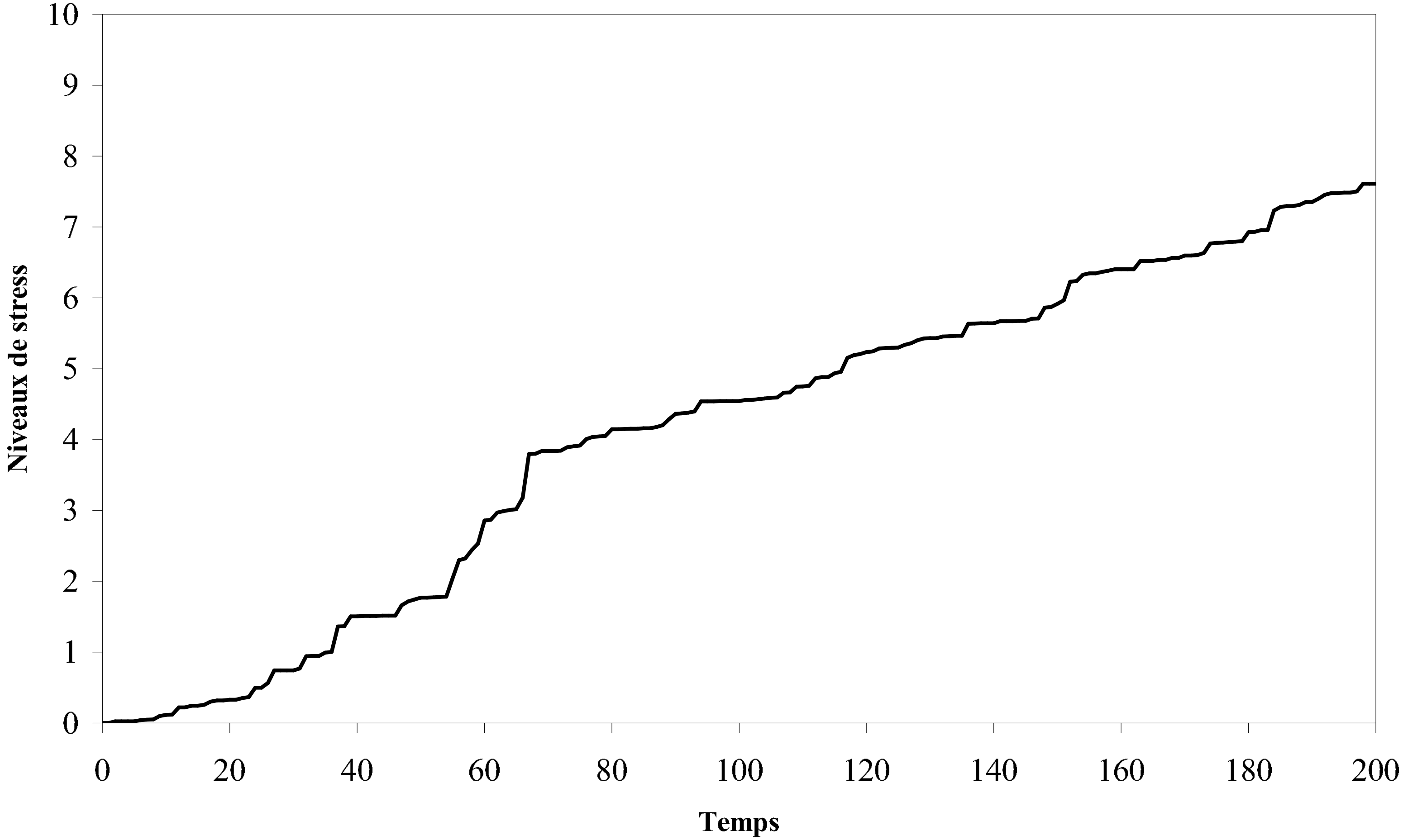
Stress x is an arbitrary step stress, $x \in E_5$.



Continuous Cyclic



Cyclic stress of type switch-on, switch-off



Degradation Stress

Notations for dynamic regression models

The **survival**, the **hazard rate** and the **cumulative hazard** and **distribution functions** given $x(\cdot)$ are :

$$S_{x(\cdot)}(t) = \mathbf{P}\left(T_{x(\cdot)} \geq t \mid x(s); 0 \leq s \leq t\right), \quad \lambda_{x(\cdot)}(t) = -\frac{S'_{x(\cdot)}(t)}{S_{x(\cdot)}(t)},$$

$$\Lambda_{x(\cdot)}(t) = -\ln\left[S_{x(\cdot)}(t)\right], \quad F_{x(\cdot)}(t) = 1 - S_{x(\cdot)}(t), \quad x(\cdot) \in E,$$

from where one can see their dependence on the **life-history** up to time t. Denote by $\lambda_0(t)$ a **baseline rate function**. It corresponds to the **mortality rate** of a given population when hypothetically all individuals are in the same **ideal, normal, usual** conditions, given by a stress $x_0(\cdot)$. We shall write also S_0 and λ_0 instead S_{x_0} and λ_{x_0}. Often we put put $x_0(\cdot) = x \equiv 0$.

On any family E of **admissible stresses**, we may consider a classe of survival functions

$$\{S_{x(\cdot)}, x(\cdot) \in E\},$$

which could be very rich.

Let $z(\cdot), y(\cdot)$ two covariates from E. We say that a **stress** $z(\cdot)$ is **accelerated with respect to a stress** $y(\cdot)$, if

$$S_{z(\cdot)}(t) \leq S_{y(\cdot)}(t), \quad t \geq 0, \quad S_{z(\cdot)}, S_{y(\cdot)} \in \{S_{x(\cdot)}, x(\cdot) \in E\}.$$

It means that if $z(\cdot)$ is an **accelerated stress with respect to a stress** $y(\cdot)$, then

$$F_{z(\cdot)}(t) \geq F_{y(\cdot)}(t), \quad t \geq 0.$$

Sometimes we write also, that

$$z(\cdot) > y(\cdot).$$

Remark on the Modelling in ALT

In reliability, **Accelerated Life Testing** in particular, the **choice of a good model** is **much more important** than in **survival analysis**. For example, in accelerated life testing units are tested under **accelerated stresses** which shorten the life. Using such experiments the life under the usual stress is estimated using some regression model. **The values of the usual stress are often not in range of the values of accelerated stresses**, since the **wide** separation between **experimental** and **usual** stresses is possible, so if the model is misspecified, the estimators of survival under the usual stress may be **very bad**.

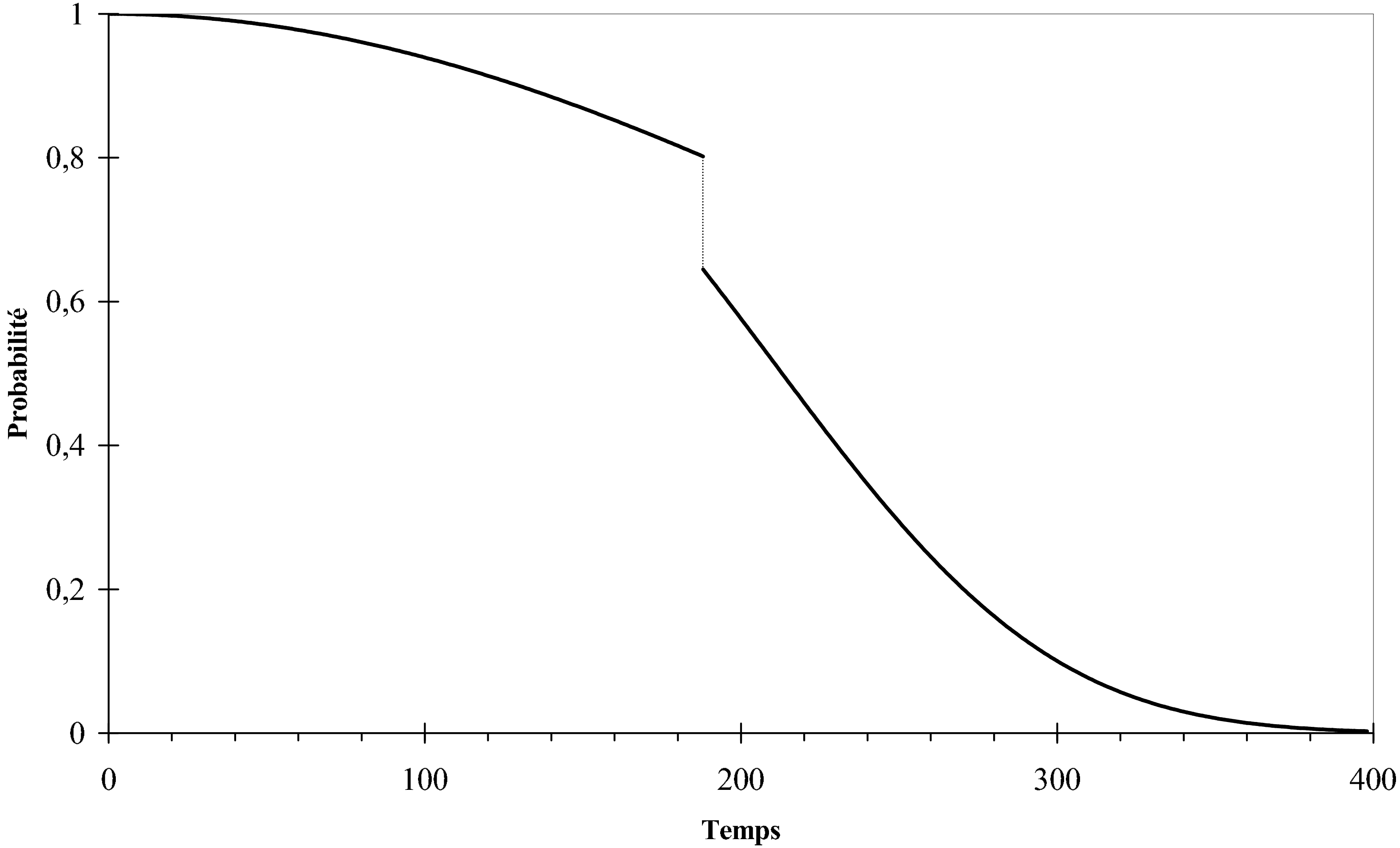
4. Sedyakin Principle and AFT model

The **physical principle in reliability** states that, for two identical populations of units functioning under stresses $x_1 \neq x_2$, two moments t_1 and t_1^* are **equivalent** if the probabilities of survival are **equal** until these moments, i.e.

$$S_{x_1}(t_1) = S_{x_2}(t_1^*).$$

This **principle, proposed by N. Sedyakin** in (1966), gives an interesting way to prolong any class of survival functions $\{S_x(\cdot), x \in E_1\}$ **indexed by constant in time stresses** to a class of survival functions indexed by step-stresses, for example from E_2 :

$$x(t) = x_1 \mathbf{1}_{\{0 \leq t < t_1\}} + x_2 \mathbf{1}_{\{t_1 \leq t\}}, \quad x_1, x_2 \in E_1.$$



Fonction de fiabilité avec une discontinuité au switch-up

According to Sedyakin we may consider the next model on E_2:

$$\lambda_{x(\cdot)}(t_1 + s) = \lambda_{x_2}(t_1^* + s), \quad \forall s \geq 0. \tag{8}$$

The meaning of this **rule of time-shift** for these step-stresses on E_2 one can see also in terms of the survival function $S_{x(\cdot)}(t),\ x(\cdot) \in E_2$ that satisfies the same **rule of time-shift**

$$S_{x(\cdot)}(t) = \begin{cases} S_{x_1}(t), & 0 \leq t < t_1, \\ S_{x_2}(t - t_1 + t_1^*), & t \geq t_1, \end{cases} \tag{9}$$

where the moment t_1^* is determined by the equality

$$S_{x_1}(t_1) = S_{x_2}(t_1^*).$$

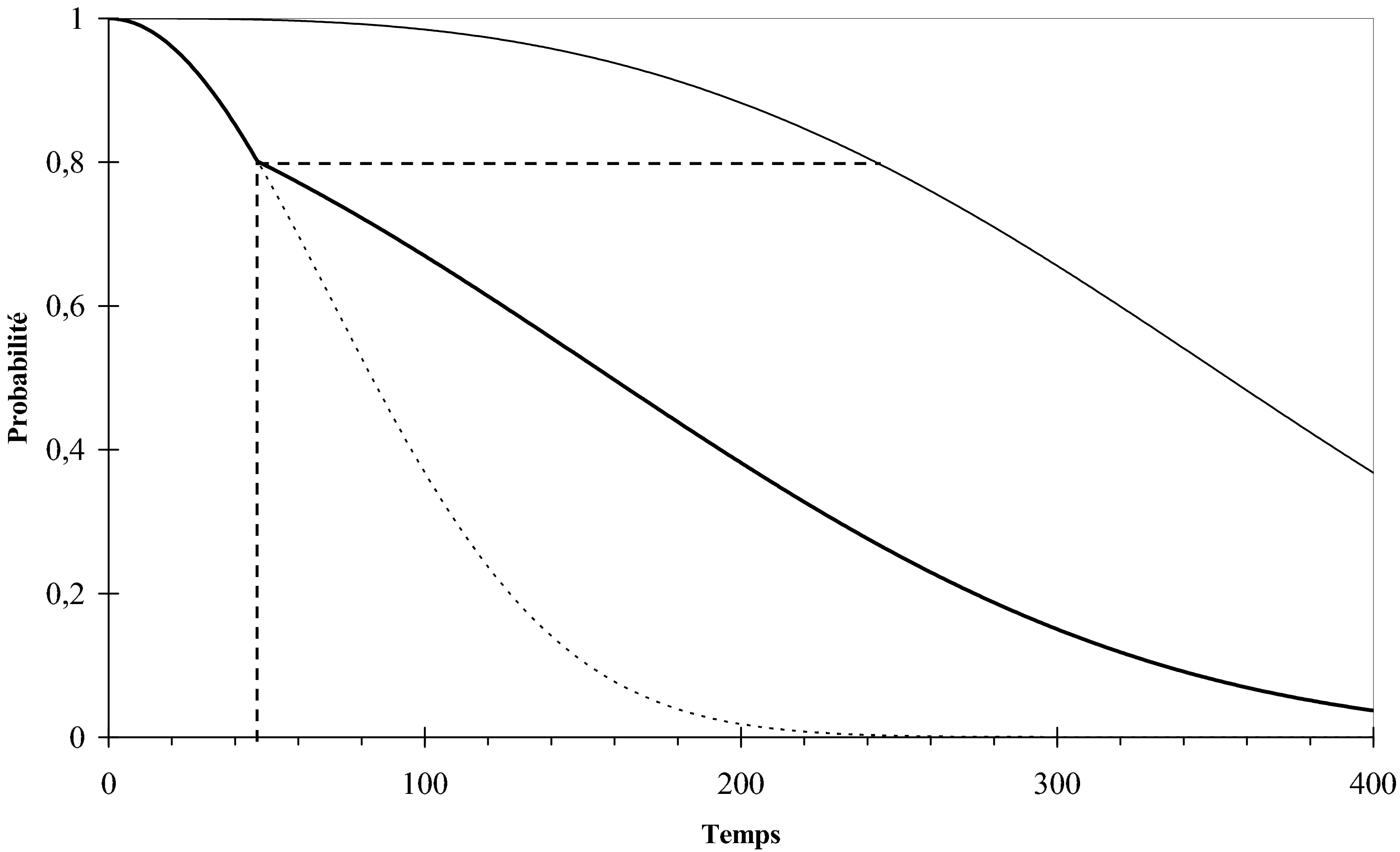
Following to L.Bolshev (1976) we call it **Sedyakin's model** on E_2.

The **GS** model generalizes this idea, by supposing that the hazard rate at any moment t depends on the stress at this moment and on the probability of survival until this moment :

$$\lambda_{x(\cdot)}(t) = g\left(x(t), S_{x(\cdot)}(t)\right), \quad x(\cdot) \in E.$$

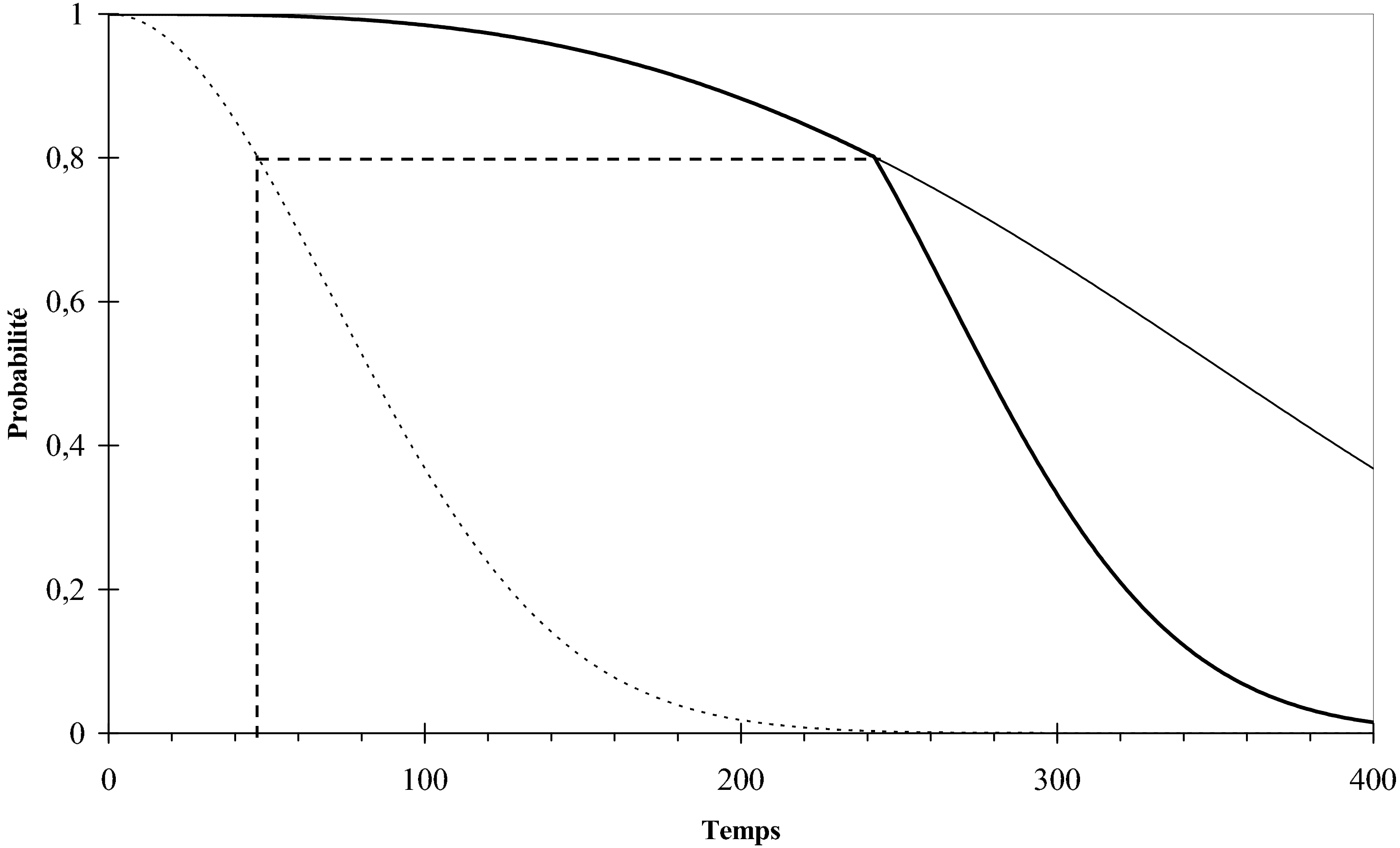
Note that the **AFT** model is the only model considered here, which verifies this rule.

The GS model is too **wide** for **AFT** data analysis but is useful for construction of narrower models. Nevertheless if the covariates are step functions, g can be more concrete.



Fonction de fiabilité sous le principe de Sedyakine

sur E_2 pour $x_1 > x_2$: endommagement préalable.



Fonction de fiabilité sous le principe de Sedyakine

sur E_2 pour $x_1 < x_2$.

Accelerated Failure Time Model

Now we consider the famous **accelerated failure time** (**AFT**) model, which is more adapted to study the aging populations (see Miner (1945), Bolshev (1976), Bagdonavičius (1978), Nelson (1990), Meeker and Escobar (1998), Viertl (1988), Bagdonavičius and Nikulin (1995, 2000, 2002), etc). The **AFT** model is the most used in **ALT**. **AFT** model holds on E if there exists a positive function $r : E \to \mathbf{R}^1$ such that for any $x(\cdot) \in E$ the survival function is given by formula:

$$S_{x(\cdot)}(t) = S_{x_0(\cdot)} \left(\int_0^t r\left[x(s)\right] ds \right) \quad , \quad x(\cdot) \in E,$$

where $x_0(\cdot)$ is a usual stress, $S_{x_0(\cdot)}$ is the so-called **baseline survival function**. The function r **changes locally** the **time-scale**.

From it follows that for **constant in time stresses** we have the next model on E_1:

$$S_x(t) = S_{x_0}(r(x)t), \quad x \in E_1.$$

Instead of $S_{x_0(\cdot)}$ we may take **any survival function** G such that the inverse $H = G^{-1}$ exists. In this case we obtain the **AFT** model on E

$$S_{x(\cdot)}(t) = G\left(\int_0^t r\,[x(s)]\,ds\right), \quad x(\cdot) \in E,$$

or as the model on E_1:

$$S_x(t) = G(r(x)t), \quad x \in E_1.$$

The **AFT** model on E can be presented in terms of **hazard rate** function by the next formula:

$$\lambda_{x(\cdot)}(t) = r\left(x(t)\right) q\left(S_{x(\cdot)}(t)\right), \quad x(\cdot) \in E,$$

for some positive q, from where it follows that it is the **Sedyakin model**. Generally **accelerated life testing experiments** are done under a stress from $E_1, (m = 1)$, sometimes under a stress from $E_2, (m = 2)$. Since **AFT** model verifies the **Sedyakin Principle**, it verifies the **rule of time-shift** E_2, according to which the survival function $S_{x(\cdot)}(t),\ x(\cdot) \in E_2$ satisfies the relation

$$S_{x(\cdot)}(t) = \begin{cases} S_{x_1}(t), & 0 \leq t < t_1, \\ S_{x_2}(t + t_1^* - t_1), & t \geq t_1, \end{cases}$$

where the moment t_1^* is determined by the equality

$$S_{x_1}(t_1) = S_{x_2}(t_1^*).$$

from where it follows that $t_1^* = \frac{r(x_1)}{r(x_2)} t_1$.

If the baseline survival function S_0 (or G) also belongs to a **parametric family**, in this case we obtain the **parametric AFT model**. For example, one may suppose that S_0 belongs to the **Power Generalized Weibull** (**PGW**) family of distributions (see, Bagdonavicius and Nikulin (2002)). In terms of the survival functions the **PGW** family is given by the next formula:

$$S(t, \sigma, \nu, \gamma) = exp\left\{1 - \left[1 + \left(\frac{t}{\sigma}\right)^{\nu}\right]^{\frac{1}{\gamma}}\right\}, t > 0, \gamma > 0, \nu > 0, \sigma > 0.$$

If $\gamma = 1$ we have the **Weibull family** of distributions. If $\gamma = 1$ and $\nu = 1 = 1$, we have the **exponential family** of distributions. This class of distributions has very nice probability properties. All moments of this distribution are finite.

In dependence of parameter values the hazard rate can be **constant, monotone** (increasing or decreasing), **unimodal** or $\bigcap$-shaped, and **bathtube** or $\bigcup$-shaped. Another interesting family, the **Exponentiated Weibull Family** of distributions, was proposed by Mudholkar & Srivastava (1995).

In practice often a baseline survival function S_0 is taken from some class of simple parametric distributions, such as **Generalized Weibull, Weibull, lognormal, log-logistique**, etc. **Parametric models** was studied by many people, see, for example, Nelson (1990), Bagdonavičius, Gerville-Réache, Nikoulina and Nikulin (2000), Bagdonavičius, Gerville-Réache and Nikulin (2002), Meeker & Escobar (1998), Sethuraman & Singpurwalla (1982), Shaked & Singpurwalla (1983), Viertl (1988), Xie (2000), Xie, Lai and Murthy (2003), etc...

Remark on the application on the AFT model

If the functions $r(\cdot)$ and $S_0(\cdot) = S_{x_0}(\cdot)$ are unknown we have a **nonparametric AFT model**. The function $r(\cdot)$ can be parametrized. If the baseline function S_0 is completely unknown, in this case we obtain a **semiparametric AFT model**.
Nonparametric and semiparametric analysis of **AFT** model was considered, for example, by Lin & Ying (1995), Duchesne & Lawless (2000,2002) , Bagdonavičius and Nikulin (2002, 2004). **AFT** model is **popular** in reliability theory because of its **interpretability**, its nice mathematical properties and its consistency with some engineering and physical principles. Nevertheless, the assumption that the survival distributions under different covariate values **differ** only in **scale** is rather **restrictive**.

5. Statistical Analysis of a redundant system with one stand-by unit

Suppose that the **failure time** T_1 of the **main element** has the **cumulative distribution function** F_1 and the **probability density** f_1, the **failure time** T_2 of the **stand-by element** has the **cumulative distribution function** F_2 and the **probability density** f_2.

The **failure time of the system** is given in terms of statistics

$$T = max(T_1, T_2).$$

Denote by $f_2^{(y)}(x)$ the **conditional density** of the **stand-by unit** given that the **main unit fails** at the moment y. It is clear that

$$f_2^{(y)}(x) = f_2(x), \quad \text{if} \quad 0 \leq x \leq y.$$

The **cumulative distribution function** $F(\cdot)$ of the **system failure time** T is given by formula

$$F(t) = P(T_1 \leq t, T_2 \leq t) = \int_0^t \left\{ \int_0^y f_2(x)dx + \int_y^t f_2^{(y)}(x)dx \right\} f_1(y)dy. \tag{1}$$

When **stand-by** is "**cold**" then

$$f_2(x) = 0, \quad f_2^{(y)}(x) = f_1(x-y), \quad x > y,$$

so

$$F(t) = \int_0^t \left\{ \int_y^t f_1(x-y)dx \right\} f_1(y)dy = \int_0^t F_1(t-y)dF_1(y).$$

In the case of "**hot**" **stand-by**

$$f_2^{(y)}(x) = f_2(x) = f_1(x), \quad \text{so} \quad F(t) = [F_1(t)]^2.$$

In the case of **”warm” reserving** the following **hypothesis is assumed**:

$$H_0 : f_2^{(y)}(x) = f_1(x + g(y) - y), \quad \text{for all} \quad x \geq y \geq 0, \tag{2}$$

where $g(y)$ is the moment which in ”**hot**” conditions **corresponds to the moment** y in ”**warm**” conditions in the **sense** that

$$F_1(g(y)) = P(T_1 \leq g(y)) = P(T_2 \leq y) = F_2(y),$$

so

$$g(y) = F_1^{-1}(F_2(y)).$$

Conditionally (given $T_1 = y$) the hypothesis corresponds to the well known **Sedyakin’s model**, Sedyakin (1966).

In Bagdonavicius and Nikoulina [1997) a goodness-of-fit test of **logrank-type** for Sediakin's model using experiments with **fixed switch off moments** is proposed (see also Bagdonavicius and Nikulin (2002)). In the situation considered here the **switch off moments** are **random**.

The formula (1) implies that **under the hypothesis** H_0

$$F(t) = \int_0^t F_1(t + g(y) - y) dF_1(y). \qquad (3)$$

In particular, if we suppose that the distribution of the units functioning in "**warm**" and "**hot**" conditions **differ only in scale**, i.e. we have

$$F_2(t) = F_1(rt), \qquad (4)$$

for some $r > 0$, then $g(y) = ry$. In such a case the hypothesis

$$H_0^* : \exists\, r > 0 : \quad f_2^{(y)}(x) = f_1(x + ry - y), \quad \text{for all} \quad x \geq y \geq 0, \qquad (5)$$

is to be verified. Conditionally (given $T_1 = y$) the hypothesis corresponds to the **accelerated failure time (AFT) model** Bolshev (1976), Nelson (1980), Bagdonavicius (1978), Bagdonavicius and Nikulin (2002).

In Bagdonavicius and Nikulin (2002) a **goodness-of-fit test** for **AFT model** using experiments with **fixed switch off moments** is proposed.

The tests

Suppose that the following **data** are available :

a) the failure times $T_{11}, \ldots, T_{1n_1}$ of n_1 units tested in **"hot" conditions**;

b) the failure times $T_{21}, \ldots, T_{2n_2}$ of n_2 units tested **in "warm" conditions**;

c) the failure times $T_1, \ldots, T_n$ of n redundant systems (**with "warm" stand-by units**).

The **tests** are based on the **difference of two estimators** of the **cumulative distribution function** $F(\cdot)$ of the **system failure time** T. The **first estimator** is the **empirical distribution function**

$$\hat{F}^{(1)}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{T_i \leq t\}}. \tag{6}$$

The **second estimator** is based on the formula (3), i.e. on the **statistics**

$$\hat{F}^{(2)}(t) = \int_0^t \hat{F}_1(t + \hat{g}(y) - y) d\hat{F}_1(y),$$

where (**hypothesis** H_0)

$$\hat{g}(y) = \hat{F}_1^{-1}(\hat{F}_2(y)), \quad \hat{F}_1^{-1}(y) = \inf\{s : \hat{F}_1(s) \geq t\}, \tag{7}$$

and

$$\hat{F}_j(t) = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{1}_{\{T_{ji} \leq t\}}, \quad j = 1, 2,$$

are the **empirical distribution functions**, or (**hypothesis** H_0^*)

$$\hat{g}(y) = \hat{r}y, \quad \hat{r} = \frac{\hat{\mu}_1}{\hat{\mu}_2}, \quad \hat{\mu}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} T_{ji}, \tag{8}$$

where μ_1 and μ_2 are the **means of two empirical distributions**.

The **test** is based on the statistic

$$X = \sqrt{n} \int_0^\infty (\hat{F}^{(1)}(t) - \hat{F}^{(2)}(t))dt. \qquad (9)$$

It is natural generalization of **Student's type** t-test for comparing the means of two populations. Indeed, the mean failure time of the system with c.d.f. F is

$$\mu = \int_0^\infty [1 - F(s)]ds,$$

so the statistic (9) is the normed difference of two estimators (the second being not the empirical mean) of the mean μ. **Student's type t-test** is based on the difference of **empirical means** of two populations.

It will be shown that in the case of both hypothesis H_0 and H_0^* the limit distribution (as $n_i/n \to l_i \in (0,1)$, $n \to \infty$) of the statistic X is normal with zero mean and finite variance σ^2.

The **test statistic** is

$$T = \frac{X}{\hat{\sigma}},$$

where $\hat{\sigma}$ is a **consistent estimator** of σ. The distribution of the statistic T **is approximated** by the **standard normal distribution** and the hypothesis H_0 (or H_0^*) **is rejected** with approximative significance value α if $\mid T \mid > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the **standard normal distribution**.

1 Limit distribution of the test statistics

Let us find the asymptotic distribution of the statistic (9).

Theorem 1. Suppose that $n_i/n \to l_i \in (0,1)$, $n \to \infty$ and the densities $f_i(x)$, $i = 1, 2$ are continuous and positive on $(0, \infty)$. Then under H_0^* the **statistic** X converges in distribution to the **normal law** $N(0, \sigma^2)$, where

$$\sigma^2 = \mathbf{Var}(T_i) + \frac{1}{l_1}\mathbf{Var}(H(T_{1i})) + \frac{c^2 r^2}{l_2}\mathbf{Var}(T_{2i}), \qquad (11)$$

where

$$H(x) = x[c+r-1-F_1(x/r)-rF_2(x)]+r\mathbf{E}(\mathbf{1}_{\{T_{1i}\leq x/r\}}T_{1i})+r\mathbf{E}(\mathbf{1}_{\{T_{2i}\leq x\}}T_{2i}),$$

$$c = \frac{1}{\mu_2}\int_0^\infty y[1 - F_2(y)]dF_1(y).$$

From the **Theorem 1** it follows that

a **consistent estimator** of the **variance** σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(T_i - \hat{\mu})^2 + \frac{n}{n_1^2}\sum_{i=1}^{n_1}[\hat{H}(T_{1i}) - \hat{\bar{H}}]^2 + \frac{\hat{c}^2\hat{r}^2 n}{n_2^2}\sum_{i=1}^{n_2}(T_{2i} - \hat{\mu}_2)^2,$$

where

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} T_i, \quad \hat{c} = \frac{1}{\hat{\mu}_2}\int_0^\infty y[1-\hat{F}_2(y)]d\hat{F}_1(y) = \frac{1}{\hat{\mu}_2}\sum_{i=1}^{n_1} T_{1i}[1-\hat{F}_2(T_{1i})],$$

$$\hat{H}(x) = x[\hat{c} + \hat{r} - 1 - \hat{F}_1(x/\hat{r}) - \hat{r}\hat{F}_2(x)] + \frac{\hat{r}}{n_1}\sum_{i=1}^{n_1}\mathbf{1}_{\{T_{1i}\le x/\hat{r}\}}T_{1i} +$$

$$\frac{\hat{r}}{n_2}\sum_{i=1}^{n_2}\mathbf{1}_{\{T_{2i}\le x\}}T_{2i},$$

$$\hat{\hat{H}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{H}(T_{1i}).$$

Theorem 2. Suppose that $n_i/n \to l_i \in (0,1)$, $n \to \infty$ and the densities $f_i(x)$, $i = 1,2$ are continuous and positive on $(0,\infty)$. Then under H_0 the **statistic** X converges in distribution to the **normal law** $N(0,\sigma^2)$, where

$$\sigma^2 = \mathbf{Var}(T_i) + \frac{1}{l_1}\mathbf{Var}(H(T_{1i})) + \frac{1}{l_2}\mathbf{Var}(Q(T_{2i}))$$

where

$$H(x) = Q(x) - xF_1(g^{-1}(x)) + g(x)[1 - F_2(x)] + \mathbf{E}(\mathbf{1}_{\{g(T_{1i}) \le x\}} g(T_{1i})) +$$

$$\mathbf{E}(\mathbf{1}_{\{T_{2i} \le x\}} g(T_{2i})) - x,$$

$$Q(x) = \mathbf{E}\{\mathbf{1}_{\{T_{1i} \le x\}}[1 - F_2(T_{1i})]/f_1(g(T_{1i}))\}.$$

A **consistent estimator** of the **variance** σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(T_i - \hat{\mu})^2 + \frac{n}{n_1^2}\sum_{i=1}^{n_1}[\hat{H}(T_{1i}) - \bar{\hat{H}}]^2 + \frac{n}{n_2^2}\sum_{i=1}^{n_2}[\hat{Q}(T_{2i}) - \bar{\hat{Q}}]^2,$$

where

$$\hat{H}(x) = \hat{Q}(x) - x\hat{F}_1(\hat{g}^{-1}(x)) + \hat{g}(x)[1 - \hat{F}_2(x)] + \frac{1}{n_1}\sum_{i=1}^{n_1}\mathbf{1}_{\{\hat{g}(T_{1i})\leq x\}}\hat{g}(T_{1i}) +$$

$$\frac{1}{n_2}\sum_{i=1}^{n_2}\mathbf{1}_{\{T_{2i}\leq x\}}\hat{g}(T_{2i}) - x, \quad \hat{Q}(x) = \frac{1}{n_1}\sum_{i=1}^{n_1}\mathbf{1}_{\{T_{1i}\leq x\}}[1 - \hat{F}_2(T_{1i})]/\hat{f}_1(\hat{g}(T_{1i})),$$

$$\hat{g}^{-1}(x) = \hat{F}_2^{-1}(\hat{F}_1(x)), \quad \hat{\hat{H}} = \frac{1}{n_1}\sum_{i=1}^{n_1} \hat{H}(T_{1i}), \quad \hat{\hat{Q}} = \frac{1}{n_2}\sum_{i=1}^{n_2} \hat{Q}(T_{2i}),$$

the **density f_1 is estimated** by the **kernel estimator**

$$\hat{f}_1(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h} K\left(\frac{x - X_{1i}}{h}\right).$$

Bibliography

1. Aalen, O., Gjessing, H. (2001). Understanding the shape of the hazard rate: a process point of view. *Statist. Sciences*, **16**, 1-22.

2. Bagdonavičius, V. (1978) Testing the hyphothesis of the additive accumulation of damages. *Probab. Theory and its Appl.*, **23**, No. 2, 403-408.

3. Bagdonavičius, V. and Nikoulina, V. (1997) A goodness-of-fit test for Sedyakin's model. *Revue Roumaine de Mathématiques Pures et Appliquées*, **42**, # 1, 5-14.

4. Bagdonavicius, V. and Nikulin, M. (1997). Some rank Tests for Multivariate Censored Data. *Advances in the Theory and Practice of Statis-*

tics, Volume in Honor of Samuel Kotz., J.Wiley and Sons, ed., 193–208.

5. Bagdonavičius V, Gerville-Réache L, Nikulin,M. (2002). On parametric inference for step-stresses models. IEEE Transaction on Reliability, # 1, 27-31.

6. Bagdonavičius V, Gerville-Réache L, Nikoulina V, Nikulin M. (2000). Expériences accélérées: analyse statistique du modèle standard de vie accélérée. Revue de Statistique Appliquée, #3, 5-38.

7. Bagdonavičius, V., Nikulin, M. (2002). *Accelerated life models.* Chapman&Hall/CRC: Boca Raton.

8. Ceci, C., Mazliak, L. (2004). Optimal Design in Nonparametric Life

Testing. *Statistical Inference for Stochastical Processes*, **7**, 305-325.

9. Cox D, Oakes D. Analysis of Survival Data. Methuen: New York, 1984.

10. Dabrowska, D. (2005). Quantile regression in transformation models. Sankhya, **67**, 153-187.

11. Dabrowska, D. (2006). Information bounds and efficient estimation in a class of censored transformation models. Webpage: arXiv:math.ST/0608088 v1 3 Aug 2006.

12. Dabrowska, D. (2006). *Estimation in a class of semi-parametric trasformation models.* In: *Second Eric L.Lehmann Symposium - Optimality.* (ed. J.Rojo), Institute of Math. Statistics, Lecture Notes and Mono-

graph Series, **49**, 166-216.

13. Duchesne, T., Lawless, J.(2000) Alternative time scale and failure time models. *Lifetime Data Analysis*,8, 263-276.

14. Duchesne, T., Rosenthal, J.S. (2003). On the collapsibility of lifetime regression models. *Advances in Applied Probabilit*,**35**, 755-772.

15. Duchesne T. Methods Common to Reliability and Survival Analysis. Recent Advances in in Reliability Theory, Limnios N, Nikulin M (eds). Birkhauser: Boston, 2000; 279-290.

16. Huber, C., Solev, V., Vonta, I. (2005). *Estimation of Density for Arbitrary Censored and Truncated Data.* Preprinte 01-2005, Mathematical

Statistics, University Victor Segalen Bordeaux 2.

17. Huber, C., Limnios, N., Mesbah, M., Nikulin, M. (2007). *Mathematical Methods in Reliability, Survival Analysis and Quality of Life*, Hermes: London.

18. Kahle, W., Wendt, H. (2006). Statistical Analysis of some parametric degradation models. In: *Probability, Statistics and Modelling in Public Health*. (Eds. M.Nikulin, D.Commenges, C.Huber). Springer, 266-279.

19. Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data.* Wiley: New York.

20. Lehmann, A. (2004). On a Degradation-Failure Models for Repairable

Items. In : *Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis, and Quality of Life*, (Eds. M.Nikulin, N.Balakrishnan, M.Mesbah, N.Limnios). Birkhauser, Boston, 65-80.

21. Limnios, N., Nikulin,M. (Eds.) (2000). *Recent Advances in Reliability Theory*, Boston: Birkhauser.

22. Lin, D.Y. and Ying, Z. (1995). Semiparametric inference for accelerated life model with time dependent covariates. *Journal of Statist. Planning and Inference*, **44**, 47-63.

23. Lindqvist, B., Doksum, K. (2003). *Mathematical and Statistical Methods in Reliability*, World Scientific: New York.

24. Martinussen, T., Scheike,T. (2006). *Dynamic regression models for survival functions.* Springer: New York.

25. Meeker, W.Q. and Escobar L., (1998), *Statistical Methods for Reliability Data*, J.Wiley and Sons, ed.

26. Mudholkar, G., Srivastava, D. (1995). The exponentiated Weibull Family: A Reanalysis of the Bus-Motor-Failure Data. *Technometrics*, **37**, #14, 436-445.

27. Nelson, W. (1990) *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses.* New York: John Wiley and Sons.

28. Nikulin, M., Commenges, D., Huber, C. (Eds.) (2006). *Probability,*

Statsitics and Modelling in Public Health, Springer: New York.

29. Nikulin, M., Gerville-Réache, Couallier, V. (2007). *Statistique des essais accélérés*. Hermes: London. #1,27-31.

30. Nikulin,M., Balakrishnan, N., Limnios,N., Mesbah, M. (Eds.) (2004). *Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis and Quality of Life*, Birkhauser: Boston.

31. Sedyakin,N.M. (1966) On one physical principle in reliability theory.(in russian). *Techn. Cybernetics*, **3**, 80-87.

32. Singpurwalla, N. (1971). Inference from accelerated life tests when observations are obtained from censored data. *Technometrics*, **13**, 161-170.

33. Veklerov, E. (1987) Reliability of redundant systems with unreliable switches. *IEEE transactions on reliability*, **36**, 470-472.

34. Viertl, R. (1988). *Statistical Methods in Accelerated Life Testing*. Goettingen: Vandenhoeck and Ruprecht.

35. Vonta, F. Nikulin, M., Limnios, Huber, C. (Eds.) (2007). *Statistical Models and Methods for Biomedical and Technical Systems*, Birkhauser: Boston.

36. Wasson Ch.S. (2005) System Analysis, Design, and Development: Concepts, Principles, and Practices. New York: Wiley.

37. Wu, H-D.I. (2004). Effect of Ignoring Heterogeneity in Hazards Re-

gression. In: *Parametric and Semipametric Models with Applicatiions to Reliability, Survival Analysis, and Quality of Life*, (eds. M.Nikulin, N.Balakrishnan, M.Mesbah, N.Limnios). Birkhauser: Boston, 239-252.

38. Yashin, A.I. (2004). Semiparametric Models in the Studies of Aging and Longevity. In: *Parametric and Semiparametric Models with Applications for Reliability, Survival Analysis, and Quality of Life*, (Eds. M.Nikulin, N.Balakrishnan, N.Limnios, M.Mesbah), Birkhauser: Boston, 149-166.