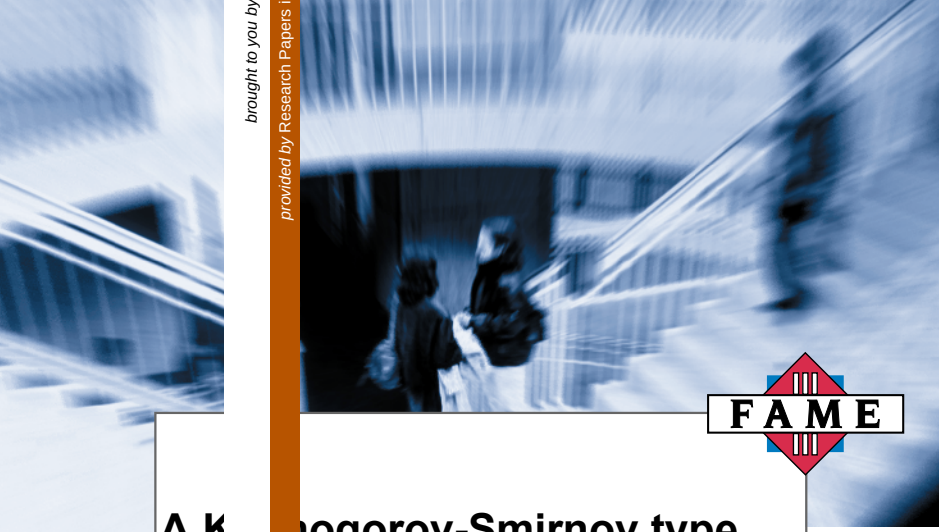
brought to you by
provided by Research Papers in
adata, citation and similar papers at core.ac.uk



FAME

A K nogorov-Smirnov type tes r shortfall dominance aga t parametric alternatives

Mic ENUIT
Institu iences Actuarielles & de Statistique, Université de Louvain

Anr cile GODERNIAUX
Institu eur Industriel Pierrard, Haute Ecole Blaise Pascal Virton

Oliv CAILLET
HEC ty of Geneva and FAME

Resea r N°143
May 2

FAME - al Center for Financial Asset Management and Engineering

UNIVERSITÉ DE GENÈVE

A Kolmogorov-Smirnov type test for shortfall dominance against parametric alternatives

Michel Denuit[1], Anne-Cécile Goderniaux[2], Olivier Scaillet[3]

April 18, 2005

Abstract

This paper proposes a Kolmogorov-type test for the shortfall order (also known in the literature as the right-spread or excess-wealth order) against parametric alternatives. In the case of the null hypothesis corresponding to the Negative Exponential distribution, this provides a test for the new better than used in expectation (NBUE) property. Such a test is particularly useful in reliability applications as well as duration and income distribution analysis. The theoretical properties of the testing procedure are established. Simulation studies reveal that the test proposed in this paper performs well, even with moderate sample sizes. Applications to real data, namely chief executive officer (CEO) compensation data and flight delay data, illustrate the empirical relevance of the techniques described in this paper.

Key words and phrases: Right-spread order, Excess-wealth order, New better than used in expectation, Bootstrap, Reliability, CEO compensation, Flight delay.

JEL Codes: C12, C14

[1] Institut des Sciences Actuarielles & Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, denuit@stat.ucl.ac.be

[2] Institut Supérieur Industriel Pierrard, Haute Ecole Blaise Pascal Virton, Belgium, anne.cecile.goderniaux@isipierrard.be

[3] HEC Genève and FAME, Université de Genève, Genève 4, Suisse, olivier.scaillet@hec.unige.ch

1 Introduction and motivation

An important aspect of reliability analysis is to find a lifetime distribution that can adequately describe the ageing behaviour of the item under study (a motor, an electronic component, a light bulb, etc.). Often, engineers are interested in the reliability of a non-repairable item (that can be anything from a small component to a large system). For such an item, the variable of interest is the time to failure or lifetime. This is the time elapsing from when the item is put into operation until it fails for the first time. For an introduction to reliability theory, we refer the reader, e.g., to Barlow & Proschan (1981) and Rausand & Hoyland (2004).

Ageing notions are used to explain how functioning items get used. Different ageing criteria have been used to classify positive and negative ageing properties. For example increasing failure rate (IFR), new better than used (NBU), decreasing mean remaining life (DMRL) and new better than used in expectation (NBUE) and their duals are the main existing ageing criteria; see Barlow & Proschan (1981). Ageing classes of life distributions are often based on comparison between survival functions of new and used items.

The following ageing notions will be encountered throughout the text. Let X be a non-negative random variable with distribution function F and survival function $\overline{F} \equiv 1 - F$. Then,

(i) X is said to be increasing failure rate (IFR, in short) if $\overline{F}$ is logconcave. It is said to be decreasing failure rate (DFR, in short) if $\overline{F}$ is logconvex. These notions correspond to the increasingness and decreasingness of the failure rate, respectively, when it exists.

(ii) X is said to be new better than used in expectation (NBUE, in short) if

$$\frac{\int_t^\infty \overline{F}(s)\,ds}{\overline{F}(t)} = \mathbb{E}[X - t|X > t] \leq \mathbb{E}[X] = \int_0^\infty \overline{F}(s)\,ds$$

for all $t \geq 0$. It is said to be new worse than used in expectation (NWUE, in short) if the reverse inequality holds for all $t \geq 0$. The quantity $\mathbb{E}[X - t|X > t]$ is called the mean residual life of the item with lifetime X: considering an item with time to failure X that is still functioning at time t, it gives the expected extra time during which the item will be working.

A machine with lifetime X that is IFR or NBUE will age with the passage of time, in the sense that its expected remaining lifetime will diminish as it gets older. On the contrary, if the lifetime X is DFR or NWUE, this means that the reliability of the machine increases as it gets older. Note that IFR implies NBUE, and that DFR implies NWUE.

Several ageing notions can also be characterized by means of a stochastic comparison with respect to the Negative Exponential distribution. We will see that the NBUE class can be obtained in this way using the shortfall order. Note further that the shortfall order and its links with the NBUE/NWUE properties also find interesting and natural applications in the analysis of duration and income distributions (see below).

The present contribution proposes a Kolmogorov-Smirnov type test for the shortfall order. References for this stochastic ordering include Shaked & Shanthikumar (1998),

Fagiuoli, Pellerey & Shaked (1999) and Fernandez-Ponce, Kochar & Muñoz-Perez (1998). As it will be pointed out in the conclusion, our approach remains nevertheless applicable for many other stochastic orderings, defined by the pointwise comparison of some transform associated to the underlying distribution functions.

The test proposed in this paper can be applied to check for the validity of the NBUE assumption. More precisely, exponentiality as a null hypothesis will be tested against the NBUE alternative. Note that other tests have been proposed in the literature. See, e.g., Klefsjo (1994) and the references therein for an overview. The testing procedure developed in the present paper will be seen to possess excellent properties (in terms of size and power).

The paper is organized as follows. Section 2 gathers some fundamental results about the shortfall order. Section 3 presents the testing procedure. In Section 4, a simulation study is provided to evaluate the finite sample performance of the test in terms of size. The power of the test is also computed against different parametric alternatives, namely Weibull, Gamma, and Pareto alternatives. Section 5 contains two empirical illustrations. First we consider compensation data for chief executive officers (CEO) of some of the largest US companies (randomly sampled from the Executive Paywatch database). It will be seen that these data exhibit NWUE, and this fact will lead to interesting interpretations about the allocation of compensations among CEO. Second we study data about European flights gathered by Eurocontrol. Both delays measured in time (minutes) and route length (kilometers) will be analyzed. In the latter case, NWUE will be detected, but not in the former. The final Section 6 concludes. Proofs are gathered in an appendix at the end of the paper.

2 Shortfall order

2.1 Expected shortfall

Let $Q(u) = \inf\{x \in \mathbb{R} | F(x) \geq u\}$ be the inverse of the distribution function F (also called the quantile function). Having a probability level u and a random variable X with distribution function F, the positive part $(X - Q(1-u))_+$ of $X - Q(1-u)$ is called the shortfall. The quantile $Q(1-u)$ is the level that is exceeded with probability u (at most). Hence, the shortfall represents the possible exceedance of X over the threshold $Q(1-u)$ exceeded by only $100u\%$ of similar devices. The expected shortfall ES is then defined as the average shortfall, that is

$$ES(u) = \mathbb{E}[(X - Q(1-u))_+] = \int_{Q(1-u)}^{+\infty} \overline{F}(x)dx.$$

The expected shortfall is an appreciated measure in reliability, finance and insurance. It possesses several nice properties, that make it appealing for practical applications. It is also closely related to the so-called excess-wealth transform defined by

$$W(u) = \frac{1}{\mathbb{E}[X]} \int_{Q(u)}^{+\infty} \overline{F}(x)dx = \frac{ES(1-u)}{\mathbb{E}[X]}, \quad u \in [0,1]. \tag{2.1}$$

In the context of economics, if X is thought of as an income, then $W(u)$ can be viewed as the proportion of the additional wealth (on top of the u-th percentile) of the richest $100(1-u)\%$ individuals in the population.

The transform W has been introduced by Kochar, Li & Shaked (2002) to compare probability distributions. Alternatively, $u \mapsto ES(1-u)$ is also called the right-spread function. This function is used to measure the spread to the right of every quantile $Q(u)$. It has been considered in Belzunce, Pinar & Ruiz (2001).

The total time on test (TTT) transform of the non-negative random variable X is defined for $u \in (0,1)$ as

$$T(u) = \mathbb{E}[X] - ES(1-u) = \int_0^{Q(u)} \overline{F}(x)dx.$$

The TTT transform is the theoretical counterpart of the empirical TTT transform that is often used in statistical reliability theory. Broadly speaking, $T(u)$ gives the average time that an item spends on test if the test is terminated when a fraction u of all the items on the test fail. For more details about the TTT transform, we refer the interested reader, e.g., to Pham & Turkkan (1994).

2.2 Shortfall order

Having two random variables X and Y, X is said to be smaller than Y in the shortfall order if the expected shortfall for X is always smaller than the corresponding expected shortfall for Y, whatever the probability level u. Specifically, denoting as F and G the distribution functions of X and Y, respectively, X precedes Y in the shortfall order, which is denoted as $X \preceq_{SH} Y$, if $ES_F(u) \leq ES_G(u)$ for all u. Considering X and Y as two times to failure, $X \preceq_{SH} Y$ means that on average, the extra time elapsed after $100u\%$ of the similar devices fail is larger for Y than for X, whatever u.

There is also a scaled version of the shortfall order. Specifically, X is said to be smaller than Y in the scaled shortfall order, denoted as $X \preceq_{SH,=} Y$ if the inequality $ES_F(u)/\mathbb{E}[X] \leq ES_G(u)/\mathbb{E}[Y]$ holds for all probability levels u. If $\mathbb{E}[X] = \mathbb{E}[Y]$ then we obviously have that $X \preceq_{SH} Y \Leftrightarrow X \preceq_{SH,=} Y$.

The scaled shortfall order is called the excess wealth order in Shaked & Shanthikumar (1998) where it is defined by means of the excess-wealth transform. The shortfall order is also termed as the right-spread order e.g., in Belzunce, Pinar & Ruiz (2001). This name originates in the fact that $\preceq_{SH}$ is based on the pointwise comparison of the right-spread functions associated with the distribution functions to be ordered.

2.3 Ageing notions and shortfall order

The classification of a lifetime according to the type of ageing structure it represents is useful, e.g., to decide about the suitability of some parametric model for the data to be analyzed. By ageing, we mean the phenomenon whereby an older system has a shorter lifetime, in some statistical sense than a younger one (after Bryson & Siddiqui (1969)). Some orderings of distributions have been used to give characterizations of ageing classes. The idea is to compare the actual distribution to the Negative Exponential distribution.

The Negative Exponential distribution is often taken as a benchmark in reliability theory. An assumption of Exponentially distributed lifetimes implies that a used item is stochastically as good as new, so there is no reason to replace a functioning item. The Negative

Exponential distribution is the most commonly used life distribution in applied reliability analysis (mainly because of its mathematical simplicity). It is therefore of interest to detect possible departures from exponentiality in the data, such as NBUE/NWUE, for instance.

Belzunce, Pinar & Ruiz (2001) proved that the class of the NBUE/NWUE distributions can be characterized with the help of the shortfall order with respect to the Negative Exponential distribution with the same mean. Specifically, given a non-negative random variable X with finite mean,

$$X \text{ NBUE} \Leftrightarrow X \preceq_{SH} \mathcal{E}xp(\mathbb{E}[X]),$$

where $\mathcal{E}xp(\mathbb{E}[X])$ represents a random variable with survival function $t \mapsto \exp(-t/\mathbb{E}[X])$. Similarly,

$$X \text{ NWUE} \Leftrightarrow \mathcal{E}xp(\mathbb{E}[X]) \preceq_{SH} X.$$

It is easy to check that if X is Exponentially distributed with mean μ, then $ES(u) = u\mu$. The NBUE property thus means that the expected shortfalls for X are smaller than the straight line with slope $\mathbb{E}[X]$, whereas the NWUE property places the expected shortfalls for X above this straight line.

3 Testing procedure

3.1 Nonparametric estimation of the expected shortfall

In this paper we consider an i.i.d. sampling scheme. Hence, we work on the basis of the following hypothesis.

Assumption 3.1. *$\{x_i\}_{i=1}^N$ is a random sample from a continuous distribution with distribution function F and mean m.*

The probability density function is denoted as usually with a lower case, namely f.

Let us introduce the empirical distribution $\hat{F}(z) := \frac{1}{N}\sum_{i=1}^N \mathbb{I}\{x_i \leq z\}$. From the theory of empirical processes we know that $\sqrt{N}(\hat{F} - F)$ converges weakly to a Brownian bridge process $\mathcal{B}_F \circ F$ (see van Der Vaart & Wellner (1986), henceforth referred to as VW).

To build the test statistic we need empirical counterparts of the moments involved in the definition of the shortfall order. Let us first deal with the nonparametric part $ES(u)/m$. The mean m will be estimated with the empirical mean:

$$\hat{m} \;:=\; \frac{1}{N}\sum_{i=1}^N x_i. \tag{3.1}$$

Note that

$$ES(u) = \int_{Q(1-u)}^{+\infty} (x - Q(1-u))dF(u) = \int_{Q(1-u)}^{+\infty} xdF(u) - uQ(1-u).$$

The expected shortfall $ES(u)$ can therefore be estimated as

$$\hat{E}S(u) \;:=\; \frac{1}{N}\sum_{i=1}^N x_i\mathbb{I}\{x_i > \hat{Q}(1-u)\} - u\hat{Q}(1-u), \tag{3.2}$$

where $\hat{Q}(1-u) := \hat{F}^{-1}(1-u)$ is the empirical quantile.

3.2 Sample distribution of the empirical expected shortfalls

Note that the estimators (3.1) and (3.2) can be viewed as particular functionals of the empirical distribution $\hat{F}$:

$$\begin{aligned} \hat{m} &= \int x d\hat{F}(x), \\ \hat{ES}(u) &= \int x \mathbb{I}\{x > \hat{F}^{-1}(1-u)\} d\hat{F}(x) - u\hat{F}^{-1}(1-u). \end{aligned}$$

To make explicit the dependence on $\hat{F}$, we will use the notation $\mathcal{J}(u;\hat{F}) := \hat{ES}(u)/\hat{m}$.

The characterisation of the test statistic in terms of a given map of the empirical distributions $\hat{F}$ is instrumental in the proof of the following lemma. The lemma describes the limiting behaviour of $\mathcal{J}(u;\hat{F})$, and is useful to deduce the properties of the testing procedure.

Lemma 3.2. *Under Assumption 3.1, $\sqrt{N}(\mathcal{J}(u;\hat{F}) - \mathcal{J}(u;F))$ converges weakly in $C((0,1))$ (the space of continuous functions on $(0,1)$) to a mean zero Gaussian process $\mathcal{J}'_F(u;\mathcal{B}_F \circ F)$ with covariance kernel given by:*

$$\Omega_F(u_1,u_2) = E[\mathcal{J}'_F(u_1;\mathcal{B}_F \circ F)\mathcal{J}'_F(u_2;\mathcal{B}_F \circ F)]$$

$$\begin{aligned} &= \mathbb{C}ov\Bigg[\frac{1}{m}(X - F^{-1}(1-u_1) + \frac{u_1}{f(F^{-1}(1-u_1))})\mathbb{I}\{X > F^{-1}(1-u_1)\} - \frac{ES(u_1)}{m^2}X, \\ &\qquad \frac{1}{m}(X - F^{-1}(1-u_2) + \frac{u_2}{f(F^{-1}(1-u_2))})\mathbb{I}\{X > F^{-1}(1-u_2)\} - \frac{ES(u_2)}{m^2}X\Bigg]. \end{aligned}$$

Remark 3.3. If we do not want to check the scaled version of the order, we can simply rely on $\bar{\mathcal{J}}(u;\hat{F}) := \hat{ES}(u)$ to build the testing procedure. A weak convergence result similar to Lemma 3.2 holds, but with a limiting process $\bar{\mathcal{J}}'_F(u;\mathcal{B}_F \circ F)$, whose covariance kernel is:

$$\begin{aligned} \bar{\Omega}_F(u_1,u_2) &= E[\bar{\mathcal{J}}'_F(u_1;\mathcal{B}_F \circ F)\bar{\mathcal{J}}'_F(u_2;\mathcal{B}_F \circ F)] \\ &= \mathbb{C}ov\Bigg[(X - F^{-1}(1-u_1) + \frac{u_1}{f(F^{-1}(1-u_1))})\mathbb{I}\{X > F^{-1}(1-u_1)\}, \\ &\qquad (X - F^{-1}(1-u_2) + \frac{u_2}{f(F^{-1}(1-u_2))})\mathbb{I}\{X > F^{-1}(1-u_2)\}\Bigg]. \end{aligned}$$

3.3 Sample distribution of the estimated parametric expected shortfalls

Let us now examine the parametric part $ES_\theta(u)/m_\theta$. An estimate $ES_{\hat{\theta}}(u)/m_{\hat{\theta}}$ obtained by plug-in may be viewed as a functional of $F_{\hat{\theta}}$, i.e. $\mathcal{J}(u;F_{\hat{\theta}})$, or as a functional of $\hat{\theta}$ itself, i.e. $J_{\hat{\theta}}(u) := \mathcal{J}(u;F_{\hat{\theta}})$. Hence we can easily characterize the limiting behavior of $\sqrt{N}(\mathcal{J}(u;F_{\hat{\theta}}) - \mathcal{J}(u;F_{\theta_0}))$ if we know the limiting behaviour of $\sqrt{N}(\hat{\theta} - \theta_0)$.

Assumption 3.4. *$\sqrt{N}(\hat{\theta} - \theta_0)$ converges to a mean zero Gaussian random variable $-\boldsymbol{V}_{\theta_0}^{-1}\boldsymbol{G}_{\theta_0}$.*

Note that this assumption is satisfied by classical M-estimators (VW Section 3.2) and Z-estimators (VW Section 3.3).

The next lemma is a direct consequence of the delta-method (VW Section 3.9), and is analogous to Lemma 3.2. Hence we omit its proof.

Lemma 3.5. *Under Assumptions 3.1 and 3.4,* $\sqrt{N}(\mathcal{J}(u;F_{\hat\theta})-\mathcal{J}(u;F_{\theta_0}))$ *converges weakly in* $C((0,1))$ *(the space of continuous functions on* $(0,1)$*) to a mean zero Gaussian process* $J'_{\theta_0}(u;-\boldsymbol{V}_{\theta_0}^{-1}\boldsymbol{G}_{\theta_0})$ *with covariance kernel given by:*

$$\Omega_{\theta_0}(u_1,u_2)=E[\dot{\boldsymbol{J}}_{\theta_0}(u_1)\boldsymbol{V}_{\theta_0}^{-1}\boldsymbol{G}_{\theta_0}\boldsymbol{G}_{\theta_0}^\top\boldsymbol{V}_{\theta_0}^{-1}\dot{\boldsymbol{J}}_{\theta_0}(u_2)^\top]$$

where

$$\dot{\boldsymbol{J}}_{\theta_0}(u)=\frac{\partial}{\partial\theta}\boldsymbol{J}_\theta(u)\Big|_{\theta=\theta_0}.$$

3.4 Testing procedure

Since we wish to test for a dominance in the $\preceq_{SH}$-sense with respect to a parametric model, namely

$$\begin{aligned} H_0:&\quad \mathcal{J}(u;F)\le\mathcal{J}(u;F_\theta)\quad \text{for all}\, u\in(0,1),\\ H_1:&\quad \mathcal{J}(u;F)>\mathcal{J}(u;F_\theta)\quad \text{for some}\, u\in(0,1),\end{aligned}$$

we consider the test statistic

$$\hat S:=\sqrt{N}\sup_u(\mathcal{J}(u;\hat F)-\mathcal{J}(u;F_{\hat\theta})),$$

and a test based on the decision rule:

$$\text{“ reject}\quad H_0\quad \text{if}\quad \hat S>c\text{”},$$

where c is some critical value that will be discussed later.

The following result characterizes the properties of the test, where

$$\bar S:=\sup_u(\mathcal{J}'_F(u;\mathcal{B}_F\circ F)-J'_{\theta_0}(u;-\boldsymbol{V}_{\theta_0}^{-1}\boldsymbol{G}_{\theta_0})).$$

Proposition 3.6. *Let* c *be a positive finite constant, then under Assumptions 3.1 and 3.4:*

i) if H_0 *is true,*

$$\lim_{N\to\infty}P[\textit{reject}\,H_0]\le P[\bar S>c]:=\alpha(c),$$

with equality when $\mathcal{J}(u;F)=\mathcal{J}(u;F_{\theta_0})$ *for all* $u\in(0,1)$*;*

ii) if H_0 *is false,*

$$\lim_{N\to\infty}P[\textit{reject}\,H_0]=1.$$

The first part of the result provides a random variable that dominates the limiting random variable corresponding to the test statistic under the null hypothesis. The inequality tells us that the test will never reject more often than $\alpha(c)$ when the null hypothesis is satisfied. Furthermore the probability of rejection will asymptotically be exactly $\alpha(c)$ when we have strict equality. The first part also implies that if one could find a c to set the $\alpha(c)$ to some desired probability level (say the conventional 0.05 or 0.01) then this would be the significance level for composite null hypotheses in the sense described by LEHMANN (1986). The second part of the result indicates that the test is capable of detecting any violation of the full set of restrictions of the null hypothesis.

Of course, in order to make the result operational, we need to find an appropriate critical value c. Since the distribution of the test statistic depends on the underlying unknown distributions, this is not an easy task, and we decide hereafter to rely on the bootstrap method to simulate p-values.

3.5 Simulating p-values

We rely on the traditional bootstrap (see BARRETT & DONALD (2003) and ABADIE (2002) for use in stochastic dominance tests). Alternatively we could use a subsampling method (as presented in POLITIS, ROMANO & WOLF (1999)) instead of a bootstrap method to get simulated p-values. The approach outlined in this section can be easily adapted to that framework.

A bootstrap sample $\{x_i^*\}_{i=1}^N$ is built from drawing N pairs with replacement from $\{x_i\}_{i=1}^N$. Let $\hat{F}^*$ denote the empirical distribution function associated to this bootstrap sample and $F_{\hat{\theta}^*}$ the distribution function associated with parametric estimation on the bootstrap sample. Let us further take

$$\hat{S}^* := \sqrt{N} \sup_u ((\mathcal{J}(u; \hat{F}^*) - \mathcal{J}(u; \hat{F})) - (\mathcal{J}(u; F_{\hat{\theta}^*}) - \mathcal{J}(u; F_{\hat{\theta}})),$$

and define

$$p^* := P[\hat{S}^* > \hat{S}].$$

Then the bootstrap method is justified by the next statement.

Proposition 3.7. *Assuming that $\alpha < 1/2$, a test for dominance in the $\preceq_{SH}$-sense based on the rule:*

$$\text{“ reject } H_0 \quad \text{if} \quad p^* < \alpha \text{”},$$

satisfies the following

$$\lim P[\text{reject } H_0] \leq \alpha \quad \text{if} \quad H_0 \text{ is true,}$$
$$\lim P[\text{reject } H_0] = 1 \quad \text{if} \quad H_0 \text{ is false.}$$

In practice we need to use Monte-Carlo methods to approximate the probability and a grid to approximate the supremum. The p-value is simply approximated by

$$p^* \approx \frac{1}{R} \sum_{r=1}^{R} \mathbb{I}\{\hat{S}_r^* > \hat{S}\},$$

where the averaging is made on R bootstrap replications and $\hat{S}_r^*$ is computed from a fine grid on $(0,1)$. Note that the replication number and the grid mesh can be chosen to make the approximations as accurate as one desires given time and computer constraints.

4 Monte Carlo results

In this section we examine the performance of the test in small and moderate samples. The replication number R to approximate the p-value is set equal to 500. A total of 250 Monte Carlo simulations are performed, and the rejection rates are computed for the bootstrap method with respect to the standard significance levels of $\alpha = 0.01$, 0.05 and 0.1.

4.1 Size

We first evaluate the Type I error (i.e. the probability to reject H_0 whereas it is true). Table 4.1 presents the results when the parametric distribution is the Negative Exponential with unit mean. In this case, the true distribution and the parametric distribution coincide. Then Proposition 3.6 suggests that the test should reject the null hypothesis H_0 with a frequency close to the chosen nominal significance level. This experiment should give us some idea about the validity of the asymptotic theory and the bootstrap method used to simulate the p-values in small samples in terms of size. The values displayed in Table 4.1 indicate that the test tends to reject the true H_0 less often than prescribed by the significance level in small samples, but the rejection rate converges to the significance level as the sample size N increases.

Table 4.1: Proportion of Negative Exponential samples where H_0 was rejected for different sample sizes with 250 simulations and 500 bootstrap replications.

Significance level	Sample size			
	N=25	N=50	N=100	$N = 200$
α=0.01	0.000	0.004	0.006	0.012
α=0.05	0.012	0.028	0.024	0.052
α=0.1	0.044	0.040	0.036	0.088

4.2 Power

4.2.1 Weibull alternative

Table 4.2 gathers the results concerning power properties with a Weibull alternative. The Weibull distribution is one of the most widely used life distribution in reliability analysis. It has been initially developed for modelling the strength of materials. Here, we use a one-parameter Weibull distribution with shape parameter less than 1, to mimick a DFR behaviour. The associated probability density function is

$$f(x) = \theta x^{\theta-1} \exp\left(-x^{\theta}\right), \quad x > 0,$$

for some positive parameter θ. Note that when $\theta = 1$ the Weibull distribution reduces to the Negative Exponential one. When $\theta = 2$, the resulting distribution is known as the Rayleigh distribution. This distribution is DFR when the shape parameter θ is less than 1, and IFR when θ is greater than 1. Therefore, the Weibull distributions are not NBUE for $\theta < 1$, and the testing procedure should therefore lead to the rejection of H_0.

We see from Table 4.2 that for small values of θ (0.25 and 0.5), the testing procedure indeed rejects H_0 for the vast majority of the samples, even for small sample sizes. The power of the test is quite high (even for small sample sizes) and increases with the sample size. For larger values of θ (i.e., when the true parent distribution is "closer" to the Exponential one), the power is weaker in small samples, but becomes reasonable in moderate samples (for a sample of size 100 and $\theta = 0.75$, the power for the usual levels of 0.05 and 0.1 are respectively equal to 72.8% and 85.2%).

Table 4.2: Proportion of samples where H_0 was rejected for different Weibull distributions and for different sample sizes with 250 simulations and 500 bootstrap replications.

Weibull parameter θ	Significance level	Sample size N=25	N=50	N=100
0.25	α=0.01	0.760	0.972	1
	α=0.05	0.952	1	1
	α=0.1	0.984	1	1
0.5	α=0.01	0.476	0.940	1
	α=0.05	0.776	0.980	1
	α=0.1	0.872	0.984	1
0.75	α=0.01	0.052	0.152	0.424
	α=0.05	0.200	0.416	0.728
	α=0.1	0.372	0.620	0.852

4.2.2 Gamma alternative

The Gamma density is also commonly used in lifetime analysis. The density function is

$$f(x) = \frac{\lambda}{\Gamma(\theta)}(\lambda x)^{\theta-1}\exp(-\lambda x), \;\; x > 0,$$

where $\theta > 0$ and $\lambda > 0$. When $\theta \geq 1$, the Gamma distributions are IFR. On the contrary, if $\theta \leq 1$ the the Gamma distributions are DFR. For $\theta = 1$, the Gamma distribution is reduced to the exponential distribution. Here, we simulate Gamma samples with $\theta < 1$ and $\lambda = 1$. These samples are not NBUE and the test procedure must reject the null hypothesis. Table 4.3 is the analogue of Table 4.2 for Gamma alternatives.

The same conclusions apply. Note that the power is weaker for Gamma distributions than for Weibull distributions. Intuitively speaking, this is due to the fact that the Gamma distribution is more similar to the Negative Exponential distribution than a Weibull distribution.

Table 4.3: Proportion of samples where H_0 was rejected for different Gamma distributions and for different sample sizes with 250 simulations and 500 bootstrap replications.

Gamma parameter θ	Significance level	Sample size N=25	N=50	N=100
0.25	α=0.01	0.476	0.900	1
	α=0.05	0.828	0.996	1
	α=0.1	0.940	1	1
0.5	α=0.01	0.096	0.300	0.756
	α=0.05	0.340	0.612	0.968
	α=0.1	0.480	0.784	0.996
0.75	α=0.01	0.016	0.018	0.056
	α=0.05	0.088	0.116	0.248
	α=0.1	0.156	0.228	0.460

4.2.3 Pareto alternative

The Pareto distribution can be obtained as a mixture of Negative Exponential distributions, or as the exponential transform of a Negative Exponentially distributed random variable (suitably shifted to have the positive half real line as support). The survival function corresponding to the one-parameter Pareto distribution is given by $\overline{F}(x) = (1+x)^{-\theta}$, $x \geq 0$, for some parameter $\theta > 0$.

The Pareto distribution always exhibits a long-tail behavior, and it is DFR for all the values of the parameter θ. Therefore, the testing procedure should reject the null hypothesis. Considering the results of the simulation study displayed in Table 4.4, we see that the power of the test increases as the sample gets bigger and as the value of θ decreases. The performance of the test is remarkable in moderate samples.

Table 4.4: Proportion of samples where H_0 was rejected for different Pareto distributions and for different sample sizes with 250 simulations and 500 bootstrap replications.

Pareto parameter θ	Significance level	Sample size N=25	N=50	N=100
1.50	α=0.01	0.436	0.700	0.912
	α=0.05	0.632	0.860	0.980
	α=0.1	0.736	0.936	0.992
2.00	α=0.01	0.268	0.444	0.780
	α=0.05	0.480	0.676	0.932
	α=0.1	0.568	0.776	0.976
2.50	α=0.01	0.200	0.316	0.572
	α=0.05	0.364	0.524	0.752
	α=0.1	0.452	0.656	0.848

5 Empirical illustrations

5.1 CEO compensation data

Compensation data for CEO of some of the largest companies in the United States are included in the Executive PayWatch database. This database is available online from http://www.aflcio.org/corporateamerica/paywatch/ceou/database.cfm. CEO compensations are listed for companies whose common stock comprises the Standard & Poor's Super 1500. We have randomly extracted 49 observations from the website database on March 8, 2005.

We have considered the total compensation, that is determined by adding the salary, bonus, other compensation, the value of restricted stock awards, long-term incentive payouts and the value of stock option awards in the fiscal year.

Figure 5.1 plots the (scaled) empirical shortfalls for the CEO compensation data (dotted line) together with the linear shortfall corresponding to the unit Negative Exponential distribution. Clearly, the empirical shortfalls dominate their Negative Exponential counterparts, except for probability levels near 1. A crossing between the two shortfall curves is visible on Figure 5.1 for probability levels higher than 0.9.

We have performed the testing procedure to decide whether this crossing invalidate the shortfall dominance of the CEO compensation parent distribution over Negative Exponential. A p-value of 4.4% is obtained (on the basis of 1000 bootstrap samples drawn from the original data set), so that the null hypothesis is rejected (at the usual level 5%).

The CEO compensation data are thus NWUE. This indicates the heavy-tail behavior of the CEO compensations. This conclusion is quite appealing. Indeed, the NWUE property ensures that $ES(u) \geq u\mu$, which allows us to write that

$$T(1-u) = \int_0^{Q(1-u)} \overline{F}(x)dx \leq (1-u)\mu.$$

Considering $T(1-u)$ as the average compensation that accrues to CEO not belonging to the top $100u\%$, this inequality indicates that these CEO earn less than $100(1-u)\%$ of the average CEO compensation. This suggests that a large part of the total compensation is concentrated on payrolls of the best paid CEO.

5.2 Delays in European flights

The data set comes from Eurocontrol. It contains all the characteristics of each flight over Europe, day by day. In this paper, we consider September 5, 2004 (but the results are similar for other days). We measure the delay first as the difference (in minutes) between the real departure time and the scheduled one, and second as the difference (in kilometers) between the real route length and the initially planned one. Here we look for a possible NWUE behavior.

Figure 5.2 displays the graphs of the (scaled) empirical expected shortfalls for the two types of delay, together with the unit Negative Exponential benchmark. For the time delay, we have randomly sampled 800 delayed flights among the 14,475 flights delayed on September 5, 2004. For the kilometer delay, we have worked with the 1,201 flights with a longer route

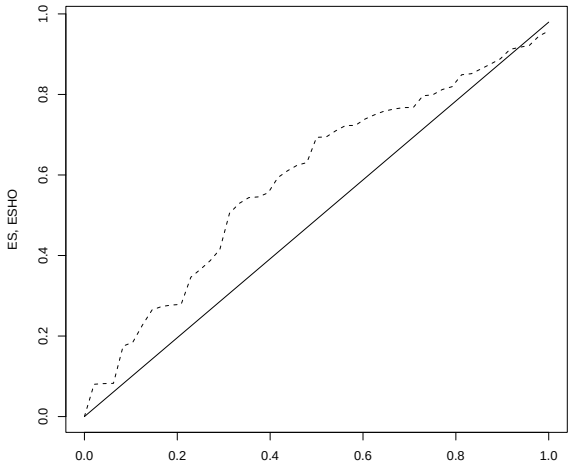


Figure 5.1: *Empirical shortfalls for the CEO compensation data set (broken line) and the corresponding linear Negative Exponential shortfalls (continuous line).*

than initially planned. In each case, we have performed 1,000 bootstrap replications of the data set. We obtain a p-value of 59% for the time delay, and of 0.4% for the length delay.

If the delay measured in minutes does not seem to exhibit a NWUE behavior, the delay measured in kilometers clearly does. Note that the single crossing between the expected shortfall curves observed on Figure 5.2 for the time delay (left panel) could also be interpreted as an ageing notion (think of an NWUE region followed by an NBUE one; see, e.g., Hawkins & Kochar (1997)).

6 Discussion

The procedure developed in this paper is applicable to any stochastic ordering defined by means of the pointwise comparison of transforms associated to the probability distributions. In particular, this applies to the moment generating function order among two non-negative random variables X and Y defined as

$$X \preceq_{MEF} Y \Leftrightarrow \mathbb{E}[\exp(tX)] \leq \mathbb{E}[\exp(tY)] \text{ for all } t > 0.$$

The order $\preceq_{MEF}$ can be tested using a Kolmogorov-Smirnov type test statistic based on empirical moment generating function functions. The theoretical properties of the empirical moment generating function process are derived in Csorgo (1982). It has been successfully applied in various testing procedures for exponentiality, e.g., by Baringhaus & Henze (1991,1992) and Henze (1993).

In this paper, only complete data are considered. Since data recorded in practice may be censored, an extension of the techniques presented in this paper to incomplete data would certainly be of interest. This issue will be addressed in a forthcoming work.

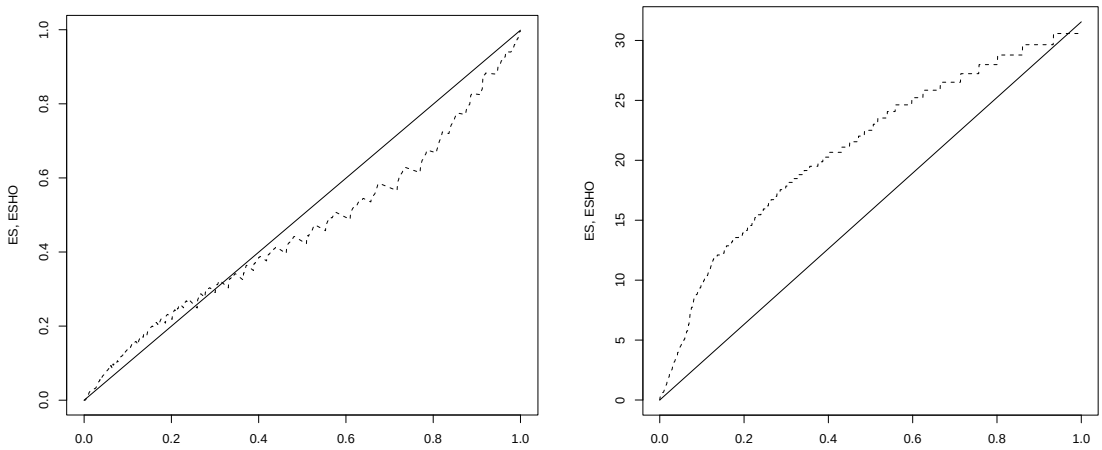


Figure 5.2: *Empirical shortfalls for the Eurocontrol data set, September 5, 2004 (broken line) for the delay measured in minutes (left panel) and in kilometers (right panel) and the corresponding linear Negative Exponential shortfalls (continuous line).*

APPENDIX

All limits are taken as N goes to infinity.

A Proof of Lemma 3.2

The result is a direct consequence of the weak convergence of the empirical process $\sqrt{N}(\hat{F} - F)$, the Hadamard differentiability of the map $\mathcal{J}$, and the delta-method (see VW Section 3.9).

B Proof of Proposition 3.6

1. Proof of Part i):

From the definitions of $\hat{S}$ and the fact that under H_0, $\mathcal{J}(u;F) - \mathcal{J}(u;F_{\theta_0}) \leq 0$ for all $u \in (0,1)$, we get that

$$\begin{aligned}\hat{S} &\leq \sup_u \sqrt{N}((\mathcal{J}(u;\hat{F}) - \mathcal{J}(u;F_{\hat{\theta}})) - (\mathcal{J}(u;F) - \mathcal{J}(u;F_{\theta_0}))) \\ &\qquad + \sup_u \sqrt{N}(\mathcal{J}(u;F) - \mathcal{J}(u;F_{\theta_0})) \\ &\leq \sup_u \sqrt{N}((\mathcal{J}(u;\hat{F}) - \mathcal{J}(u;F_{\hat{\theta}})) - (\mathcal{J}(u;F) - \mathcal{J}(u;F_{\theta_0}))).\end{aligned}$$

Hence the results follows from the weak convergence of $\sqrt{N}((\mathcal{J}(u;\hat{F}) - \mathcal{J}(u;F_{\hat{\theta}})) - (\mathcal{J}(u;F) - \mathcal{J}(u;F_{\theta_0})))$ induced by Lemmas 3.2 and 3.5, and the definition of $\bar{S}$.

2. Proof of Part *ii*):

If the alternative is true, then there is some u, say $\bar{u} \in (0,1)$, for which $\mathcal{J}(\bar{u};F) - \mathcal{J}(\bar{u};F_{\theta_0}) = \delta > 0$. Then the result follows using the inequality $\hat{S} \geq \sqrt{N}(\mathcal{J}(\bar{u};\hat{F}) - \mathcal{J}(\bar{u};F_{\hat{\theta}}))$ and almost sure uniform convergence.

C Proof of Proposition 3.7

We know that $\sqrt{N}(\hat{F}^* - \hat{F})$ converges weakly to an independent copy of $\mathcal{B}_F \circ F$ (VW Theorem 3.6.3). Then Hadamard differentiability, via the delta-method for bootstrap (VW Theorem 3.9.11), and the CMT, yield that $\hat{S}^*$ converges in probabilility to a random variable, which is an independent copy of $\bar{S}$. Note that the distribution $P^0(t)$ of this random variable is absolutely continuous (Tsirel'son (1975)), while its median is strictly positive and finite. Moreover $c(\alpha)$ defined by $P(\bar{S} > c(\alpha)) = \alpha$ is finite and positive for any $\alpha < 1/2$ (VW Proposition A.2.7).

Note that the event $\{p^* < \alpha\}$ is equivalent to the event $\{\hat{S} > \hat{c}^*(\alpha)\}$ where

$$\inf\{t : \hat{P}^*(t) > 1 - \alpha\} = \hat{c}^*(\alpha) \xrightarrow{p} c(\alpha), \tag{C.1}$$

by the convergence of $\hat{S}^*$ and the aforementioned properties of P^0. Then:

$$\begin{aligned} \lim P[\text{reject } H_0 | H_0] &= \lim P(\hat{S} > \hat{c}^*(\alpha)) \\ &= \lim P(\hat{S} > c(\alpha)) + \lim(P(\hat{S} > \hat{c}^*(\alpha)) - P(\hat{S} > c(\alpha))) \\ &\leq P(\bar{S} > c(\alpha)) := \alpha, \end{aligned}$$

where the last statement comes from (C.1), part *i*) of Proposition 3.6 and $c(\alpha)$ being a continuity point of the distribution of $\bar{S}$. On the other hand part *ii*) of Proposition 3.6 and $c(\alpha) < \infty$ ensure that $\lim P[\text{reject } H_0 | H_1] = 1$.

Aknowledgements

The authors express their gratitude to Claus Gwiggner for having provided the Eurocontrol data, as well as for providing relevant information about the contents of this data set.

Michel Denuit and Olivier Scaillet gratefully acknowledge the financial support of the Belgian Government through the "Projet d'Action de Recherches Concertées" 04/09-320, and of the Swiss National Science Foundation through the National Center of Competence in Research: Financial Valuation and Risk Management, respectively.

References

[1] Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association* **97**, 284-292.

[2] Baringhaus, L., & Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transforms. *Annals of the Institute of Statistical Mathematics* **43**, 551-564.

[3] Baringhaus, L., & Henze, N. (1992). An adaptive omnibus test for exponentiality. *Communications in Statistics: Theory and Methods* **21**, 969-978.

[4] Barlow, R.E., & Proschan, F. (1981). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, New York.

[5] Barrett, G., & Donald, S. (2003). Consistent tests for stochastic dominance. *Econometrica* **71**, 71-104.

[6] Belzunce, F., Pinar, J.F., & Ruiz, J.M. (2001). A family of tests for right spread order. *Statistics and Probability Letters* **54**, 79-92.

[7] Bryson, M.C., & Siddiqui, M.M. (1969). Some criteria for ageing. *Journal of the American Statistical Association* **64**, 1472-1482.

[8] Csorgo, S. (1982). The empirical moment generating function. In *Nonparametric Statistical Inference*, edited by B.V. Gnedenko, M.L. Puri & I. Vincze, Vol. 32, pp. 139-150, Coll. Math. Soc. J. Bolyai, Elsevier, Amsterdam.

[9] Fagiuoli, E., Pellerey, F., & Shaked, M. (1999). A characterization of the dilation order and its applications. *Statistical Papers* **40**, 393–406.

[10] Fernandez-Ponce, J. M., Kochar, S. C., & Muñoz-Perez, J. (1998). Partial orderings of distributions based on right-spread functions. *Journal of Applied Probability* **35**, 221–228.

[11] Hawkins, D.L., & Kochar, S. (1997). Inference about the transition point in NBUE-NWUE models. *Sankhya - Series A* **59**, 117-132.

[12] Henze, N. (1993). A new flexible class of omnibus tests for exponentiality. *Communications in Statistics: Theory and Methods* **22**, 115-133.

[13] Klefsjö, B. (1994). Letter to the Editor on "An exact test for NBUE class of survival functions". *Communications in Statistics: Theory and Methods* **23**, 2417-2421.

[14] Kochar, S.C., Li, X., & Shaked, M. (2002). The total time on test transform and the excess wealth stochastic orders of distributions. *Advances in Applied Probability* **34**, 826-845.

[15] Lehmann, E. (1986). *Testing Statistical Hypotheses*. Wiley & Sons, New York.

[16] Pham, T.G., & Turkkan, N. (1994). The Lorenz and the scaled total-time-on-test transform curves: a unified approach. *IEEE Transactions on Reliability* **43**, 76-84.

[17] Politis, D., Romano, J., & Wolf, M. (1999). *Subsampling*. Springer Verlag, New York.

[18] Rausand, M, & Hoyland, A. (2004). *System Reliability Theory*. Wiley, New York.

[19] Shaked, M., & Shanthikumar, J.G. (1998). Two variability orders. *Probability in the Engineering and Informational Sciences* **12**, 1–23.

[20] Tsirel'son, V. (1975). The density of the distribution of the maximum of a gaussian process. *Theory of Probability and their Applications* **16**, 847-856.

[21] van der Vaart, A., & Wellner, J. (1996). *Weak Convergence and Empirical Processes*. Springer Verlag, New York.

The FAME Research Paper Series

The International Center for Financial Asset Management and Engineering (FAME) is a private foundation created in 1996 on the initiative of 21 leading partners of the finance and technology community, together with three Universities of the Lake Geneva Region (Switzerland). FAME is about ***Research, Doctoral Training****, and* ***Executive Education*** *with "interfacing" activities such as the FAME lectures, the Research Day/Annual Meeting, and the Research Paper Series.*

The ***FAME Research Paper Series*** *includes three types of contributions: First, it reports on the research carried out at FAME by students and research fellows; second, it includes research work contributed by Swiss academics and practitioners interested in a wider dissemination of their ideas, in practitioners' circles in particular; finally, prominent international contributions of particular interest to our constituency are included on a regular basis. Papers with strong practical implications are preceded by an Executive Summary, explaining in non-technical terms the question asked, discussing its relevance and outlining the answer provided.*

Martin Hoesli is acting Head of the Research Paper Series. Please email any comments or queries to the following address: Martin.Hoesli@hec.unige.ch.

The following is a list of the 10 most recent FAME Research Papers. For a complete list, please visit our website at www.fame.ch under the heading 'Faculty and Research, Research Paper Series, Complete List'.

N°142 **Times-to-Default: Life Cycle, Global and Industry Cycle Impacts**
Fabien COUDREC, University of Geneva and FAME, Olivier RENAULT, FERC, Warwick Business School

N°141 **Understanding Default Risk Through Nonparametric Intensity Estimation**
Fabien COUDREC, University of Geneva and FAME

N°140 **Robust Mean-Variance Portfolio Selection**
Cédric PERRET-GENTIL, Union Bancaire Privée, Maria-Pia VICTORIA-FESER, HEC, University of Geneva

N°139 **Trading Volume in Dynamically Efficient Markets**
Tony BERRADA, HEC Montréal, CIRANO and CREF, Julien HUGONNIER, University of Lausanne, CIRANO and FAME, Marcel RINDISBACHER, Rotman School of Management, University of Toronto and CIRANO.

N°138 **Growth Options in General Equilibrium: Some Asset Pricing Implications**
Julien HUGONNIER, University of Lausanne and FAME, Erwan MORELLEC, University of Lausanne, FAME and CEPR, Suresh SUNDARESAN, Graduate School of Business, Columbia University.

N°137 **On the Demand for Budget Constrained Insurance**
Richard WATT, Universidad Autónoma de Madrid, Henri LOUBERGÉ, University of Geneva and FAME

N°136 **Direct Prference for Wealth in Aggregate Household Portfolios**
Pascal ST-AMOUR, HEC, University of Lausanne, FAME, CIRANO and CIRPEE

N°135 **Indirect Robust Estimation of the Short-Term Interest Rate Process**
Veronika CZELLAR, Dept. of Econometrics, University of Geneva, G. Andrew KAROLYI, Fisher College of Business, Ohio State University, Elvezio RONCHETTI, Dept. of Econometrics, University of Geneva.

N°134 **Do Major Financial Crises Provide Information on Sovereign Risk to the Rest of the World? A Look at Credit Default Swap Markets**
Didier COSSIN, IMD International and FAME, Gero JUNG, Graduate Institute of International Studies.

N°133 **Are European Corporate Bond and Default Swap Markets Segmented?**
Didier COSSIN, IMD International, Hongze LU, IMD International, HEC, University of Lausanne.

International Center FAME - Partner Institutions

The University of Geneva

The University of Geneva, originally known as the Academy of Geneva, was founded in 1559 by Jean Calvin and Theodore de Beze. In 1873, The Academy of Geneva became the University of Geneva with the creation of a medical school. The Faculty of Economic and Social Sciences was created in 1915. The university is now composed of seven faculties of science; medicine; arts; law; economic and social sciences; psychology; education, and theology. It also includes a school of translation and interpretation; an institute of architecture; seven interdisciplinary centers and six associated institutes.

More than 13'000 students, the majority being foreigners, are enrolled in the various programs from the licence to high-level doctorates. A staff of more than 2'500 persons (professors, lecturers and assistants) is dedicated to the transmission and advancement of scientific knowledge through teaching as well as fundamental and applied research. The University of Geneva has been able to preserve the ancient European tradition of an academic community located in the heart of the city. This favors not only interaction between students, but also their integration in the population and in their participation of the particularly rich artistic and cultural life. ***http://www.unige.ch***

The University of Lausanne

Founded as an academy in 1537, the University of Lausanne (UNIL) is a modern institution of higher education and advanced research. Together with the neighboring Federal Polytechnic Institute of Lausanne, it comprises vast facilities and extends its influence beyond the city and the canton into regional, national, and international spheres.

Lausanne is a comprehensive university composed of seven Schools and Faculties: religious studies; law; arts; social and political sciences; business; science and medicine. With its 9'000 students, it is a medium-sized institution able to foster contact between students and professors as well as to encourage interdisciplinary work. The five humanities faculties and the science faculty are situated on the shores of Lake Leman in the Dorigny plains, a magnificent area of forest and fields that may have inspired the landscape depicted in Brueghel the Elder's masterpiece, the Harvesters. The institutes and various centers of the School of Medicine are grouped around the hospitals in the center of Lausanne. The Institute of Biochemistry is located in Epalinges, in the northern hills overlooking the city. ***http://www.unil.ch***

The Graduate Institute of International Studies

The Graduate Institute of International Studies is a teaching and research institution devoted to the study of international relations at the graduate level. It was founded in 1927 by Professor William Rappard to contribute through scholarships to the experience of international co-operation which the establishment of the League of Nations in Geneva represented at that time. The Institute is a self-governing foundation closely connected with, but independent of, the University of Geneva.

The Institute attempts to be both international and pluridisciplinary. The subjects in its curriculum, the composition of its teaching staff and the diversity of origin of its student body, confer upon it its international character. Professors teaching at the Institute come from all regions of the world, and the approximately 650 students arrive from some 60 different countries. Its international character is further emphasized by the use of both English and French as working languages. Its pluralistic approach - which draws upon the methods of economics, history, law, and political science - reflects its aim to provide a broad approach and in-depth understanding of international relations in general. ***http://heiwww.unige.ch***

FAME

INTERNATIONAL CENTER FOR
FINANCIAL ASSET MANAGEMENT AND ENGINEERING

40, Bd. du Pont d'Arve
PO Box, 1211 Geneva 4
Switzerland
Tel (++4122) 312 09 61
Fax (++4122) 312 10 26
http: //www.fame.ch
E-mail: admin@fame.ch

UNIVERSITÉ DE GENÈVE

THE GRADUATE INSTITUTE OF
INTERNATIONAL STUDIES

UNIVERSITÉ DE LAUSANNE