Limit theorems for $M(t)/G/\infty$ queuing system with a heavy tailed distribution of servise times

Chernavskaya Ekaterina[1]

Department of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia
(E-mail: Chernavskayaak@mail.ru)

Abstract. This article focuses on queuing systems with the Poisson input flow and an infinite number of servers. Service times have the heavy-tailed distribution. We obtain conditions under which finite-dimensional distributions of the normalized process, which determines the number of occupied servers, weakly converge with increasing time to finite-dimensional distributions of the Gaussian process.

Keywords: queuing systems, the number of occupied servers, an infinite number of servers, heavy-tailed distribution, finite-dimensional distributions.

1 Introduction

In our article we regard a convergence of finite-dimensional distributions for the process of number of occupied servers $q(t)$ in a queuing system with an infinite number of servers. For such a system an extensive literature is devoted. For example, infinite server queuing systems with restrictions([7], [11]), infinite server queuing systems in a random environment [6], infinite network systems [10] and others were considered. This is due to a wide range of practical issues in which these models are useful, and a number of emerging here interesting mathematical problems.

At first glance, infinite channel systems seem to be unrealistic, but in fact, they can be considered as models of many real-world objects. For example, in the communication theory, in the study of total flow of impulses [1], as well as in the description of the formation of queues at the crossroads of unmanaged highways [2], in some problems of security(see [9] about relation between risk model of Cramer-Lundberg model and queuing system $G/G/\infty$).

In addition, these models can be considered as approximations of systems with a large number of servers. Note that approaches used for studying such systems are useful for queuing problems in the case of the high load.

When we study some complex queuing system, it is sometimes useful to consider an infinite channel system as the easiest for studying. It is also convenient to consider such systems, when in the system a queue does not form, for some reasons [12].

If service time has a finite mathematical expectation then there exists a proper limit distribution for the number of customers $q(t)$ in the system at time t as $t \to \infty$. But it is not the case when distribution function of service times has a heavy tail, i.e. there is no the mean of the service time. In this situation $q(t)$ goes to infinity as $t \to \infty$ and the problem of asymptotic analysis of it's behavior occurs.

16th ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece

We consider a system with a Poisson input flow with intensity $\lambda(t)$. Our aim is to find conditions under which finite-dimensional distributions of the normalized process, which determines the number of occupied servers, weakly converge with increasing time to finite-dimensional distributions of the Gaussian process.

2 Description of the system

We consider a queuing system S with an infinite number of servers. Arriving customers form a Poisson process $X(t)$ with the intensity $\lambda(t)$. Assume that the intensity $\lambda(t)$ of the input flow satisfies the following

Condition 1 *Denote* $\Lambda(t) = \int\limits_0^t \lambda(y)dy$. *There exist a finite* $\tau \geq 0$, $\lambda > 0$ *and the sequence* $\{S_k\}_{k=0}^{\infty}$ *such that* $S_k \to \infty$ *as* $k \to \infty$ *and*

$$0 < S_k - S_{k-1} \leq \tau \text{ and } \Lambda(S_k) = \lambda S_k,\ k \geq 1,\ S_0 = 0.$$

It should be noted that from the *Condition 1* follows that the limit of $\frac{1}{t}\Lambda(t)$ exists as $t \to \infty$ and equals to λ.

We assume that service times $\{\eta_i\}_{i=1}^{\infty}$ are independent identically distributed(i.i.d.) random variables with a distribution function $B(x)$, $\overline{B}(x) = 1 - B(x)$. Suppose that this function satisfies the following

Condition 2

$$\overline{B}(t) \sim \frac{\mathcal{L}(t)}{t^{\beta}} \text{ as } t \to \infty, \tag{1}$$

for some $0 < \beta < 1$ *and slowly varying at infinity function* $\mathcal{L}(t)$ *[5].*
For some functions $f(t)$ *and* $g(t)$ *the notation* $f(t) \sim g(t)$ *as* $t \to \infty$ *means that* $\lim\limits_{t\to\infty} \frac{f(t)}{g(t)} = 1$.

Let $q(t)$ be the number of customers in the system at time t and $q(0) = 0$ in probability. Our aim is to study the asymptotic behavior of the process $q(tT)$ as $T \to \infty$ for $t \in (0, h)$, $h > 0$. Let us formulate our main result.

Theorem 1 *Suppose that Conditions 1 and 2 are fulfilled, then finite-dimensional distributions of the process*

$$\frac{q(tT) - \rho(tT)}{\sqrt{\mathcal{L}(T)T^{1-\beta}}}$$

weakly converge as $T \to \infty$ *to finite-dimensional distributions of the centered Gaussian process* $\xi(t)$ *with the covariance function*

$$R(t, t+u) = \frac{\lambda}{1-\beta}\left((t+u)^{1-\beta} - u^{1-\beta}\right),\ (t \geq 0, u \geq 0).$$

Here $\rho(tT) = \int\limits_0^{tT} \overline{B}(Tt - x)\lambda(x)dx$, $t \in (0, h)$.

In order to clarify the basic idea which was used in the proof of the *Theorem 1*, we begin by consideration of a simple case.

3 A particular case: the intensity of the input flow is a constant

We assume that customers which enter the system form a Poisson process with the intensity λ. Note that the *Condition 1* is always satisfied in this case. We denote this queuing system by S_1.

In order to prove the *Theorem 1* let us fix some moments of time $0 < t_1 < \ldots < t_{n-1} < t_n < \infty$. We denote the number of customers which enter the system on $[Tt_{i-1}; Tt_i]$ and be served on the interval $[Tt_j; Tt_{j+1}]$ as ξ_{ij}^T for some fixed T where $1 \le i \le n$, $1 \le j \le n$. Here we assume that $t_0 = 0$, $t_{n+1} = \infty$.

Lemma 1 *For random variables ξ_{ij}^T, $i \le j$, following statements are true:*

- *ξ_{ij}^T and ξ_{kl}^T are independent for $i \ne k$, $j \ne l$;*
- *ξ_{ij}^T has a Poisson distribution with the parameter α_{ij}^T where*

$$\alpha_{ij}^T = \lambda \int_{Tt_{i-1}}^{Tt_i} \left(\overline{B}(Tt_j - y) - \overline{B}(Tt_{j+1} - y)\right) dy$$

for $1 \le i \le n$, $1 \le j \le n$, $i \le j$.

Proof. We introduce the notation

$$\boldsymbol{\xi_i^T} = (\xi_{ii}^T, \xi_{ii+1}^T, \ldots, \xi_{in}^T)$$

where $1 \le i \le n$.

Note that the independence of vectors $\boldsymbol{\xi_k}^T$ and $\boldsymbol{\xi_l}^T$ for $k \ne l$ follows from the property of independence of the jumps' number of a Poisson process on disjoint intervals and the independence of service times.

Let us show the independence of coordinates for the vector $\boldsymbol{\xi_i}^T$, $1 \le i \le n$. To do this, we fix some non-negative integers $\{k_j\}_{j=i}^n$. Denote by p_{ij}^T the probability that a customer which enters the system on the interval $[Tt_{i-1}, Tt_i]$ is served on $[Tt_j, Tt_{j+1}]$. Using the formula of the total probability we find the joint distribution of coordinates of the vector $\boldsymbol{\xi_i}^T$

$$P\left(\xi_{ii}^T = k_i, \xi_{ii+1}^T = k_{i+1}, \ldots, \xi_{in}^T = k_n\right) =$$

$$= \sum_{N \ge k_i + k_{i+1} + \ldots + k_n} \frac{N!}{k_i! k_{i+1}! \ldots k_n! (N - k_i - k_{i+1} - \ldots - k_n)!} \times$$

$$\times \left(p_{ii}^T\right)^{k_i} \left(p_{ii+1}^T\right)^{k_{i+1}} \ldots \left(p_{in}^T\right)^{k_n} \left(1 - p_{ii}^T - p_{ii+1}^T - \ldots - p_{in}^T\right)^{N - k_i - k_{i+1} - \ldots - k_n} \times$$

$$\times \frac{\left(Tt_i - Tt_{i-1}\right)^N}{N!} e^{-\lambda(Tt_i - Tt_{i-1})} =$$

$$= \prod_{j=i}^n \frac{\left(\lambda p_{ij}^T (Tt_i - Tt_{i-1})\right)^{k_j}}{k_j!} e^{-\lambda p_{ij}^T (Tt_i - Tt_{i-1})} = \prod_{j=i}^n P\left(\xi_{ij}^T = k_j\right).$$

Thus, we obtain that coordinates of a vector $\boldsymbol{\xi_i}^T$ are independent. The statement, that ξ_{ij}^T has a Poisson distribution with a parameter α_{ij}^T, follows from the chain of equalities of previous reasoning, so

$$P\left(\xi_{ij}^T = k_j\right) = \frac{\left(\lambda p_{ij}^T(Tt_i - Tt_{i-1})\right)^{k_j}}{k_j!} e^{-\lambda p_{ij}^T(Tt_i - Tt_{i-1})}$$

where $p_{ij}^T = \frac{1}{Tt_i - Tt_{i-1}} \int\limits_{Tt_{i-1}}^{Tt_i} \left(\overline{B}(Tt_j - y) - \overline{B}(Tt_{j+1} - y)\right) dy$. Thus, the second part of the lemma is proved. ■

Lemma 2 *If the Condition 2 is satisfied then for functions* α_{ij}^T, $1 \le i \le j \le n$ *the following asymptotic behavior takes place*

$$\alpha_{ij}^T \sim \frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta} d_{ij} \ \textit{as}\ T \to \infty$$

where $d_{ij} = (t_{j+1} - t_i)^{1-\beta} + (t_j - t_{i-1})^{1-\beta} - (t_j - t_i)^{1-\beta} - (t_{j+1} - t_{i-1})^{1-\beta}$.

Proof. In order to facilitate calculations we introduce the following notation

$$\mu_{ij}^T = \lambda \int_0^{Tt_i} \overline{B}(Tt_j - y) dy, \ i \le j,$$

then α_{ij}^T can be rewritten as

$$\alpha_{ij}^T = \mu_{ij}^T - \mu_{i-1j}^T - \mu_{ij+1}^T + \mu_{i-1j+1}^T.$$

Let us find the asymptotic behavior of μ_{ij}^T. For this we consider following cases.

1. Let $i < j$. It follows from the *Condition 2* that for any $\varepsilon > 0$ and sufficiently large t the following inequalities hold

$$(1-\varepsilon)\frac{\mathcal{L}(t)}{t^\beta} \le \overline{B}(t) \le \frac{\mathcal{L}(t)}{t^\beta}(1+\varepsilon). \tag{2}$$

   Similarly, from the definition of a slowly varying function it follows that for any $a > 0$, $\varepsilon > 0$ and sufficiently large t

$$(1-\varepsilon)\mathcal{L}(t) \le \mathcal{L}(at) \le \mathcal{L}(t)(1+\varepsilon). \tag{3}$$

   It follows from (2) and (3) that for any $\varepsilon > 0$ there is $\bar{t}$ such that for $t > \bar{t}$ we have

$$\mu_{ij}^T = \lambda \int_0^{Tt_i} \overline{B}(Tt_j - y) dy = \lambda T \int_0^{t_i} \overline{B}(T(t_j - z)) dz \ge$$

$$\ge (1-\varepsilon)\lambda T \int_0^{t_i} \frac{\mathcal{L}(T(t_j - z))}{(T(t_j - z))^\beta} dz = \lambda(1-\varepsilon) T^{1-\beta} \int_0^{t_i} \frac{\mathcal{L}(T(t_j - z))}{(t_j - z)^\beta} \frac{\mathcal{L}(T)}{\mathcal{L}(T)} dz \ge$$

$$\geq (1-\varepsilon)^2 \lambda \mathcal{L}(T) T^{1-\beta} \int_0^{t_i} \frac{dz}{(t_j - z)^\beta} = (1-\varepsilon)^2 \frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta} \left(t_j^{1-\beta} - (t_j - t_i)^{1-\beta} \right).$$

Similarly, using the inequalities (2) and (3), we obtain an upper bound for μ_{ij}^T. Combining these results we get the equivalence

$$\mu_{ij}^T \sim \frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta} \left(t_j^{1-\beta} - (t_j - t_i)^{1-\beta} \right) \text{ for any } 1 \leq i < j \leq n \quad (4)$$

as $T \to \infty$.

2. Secondly, let $i = j$. Denote

$$I(t) = \int_0^t \overline{B}(y) dy.$$

Then $\mu_{ii}^T = \lambda I(T t_i)$. Let us find the asymptotic behavior of the integral $I(t)$ as $t \to \infty$. For any $0 < \gamma < 1$ the following representation holds

$$I(t) = \int_0^{t^\gamma} \overline{B}(y) dy + \int_{t^\gamma}^t \overline{B}(y) dy = I_1(t) + I_2(t).$$

Thus, now we find the asymptotic behavior of $I_1(t)$ and $I_2(t)$.

- Using inequalities (2) and (3), we get that for any $\varepsilon > 0$ there is $\widetilde{t}$ such that for all $t > \widetilde{t}$

$$I_2(t) = \int_{t^\gamma}^t \overline{B}(y) dy = t \int_{t^{\gamma-1}}^1 \overline{B}(zt) dz \geq (1-\varepsilon) t \int_{t^{\gamma-1}}^1 \frac{\mathcal{L}(tz)}{(tz)^\beta} dz =$$

$$= (1-\varepsilon) t^{1-\beta} \int_{t^{\gamma-1}}^1 \frac{\mathcal{L}(tz)}{z^\beta} \frac{\mathcal{L}(t)}{\mathcal{L}(t)} dz \geq (1-\varepsilon)^2 t^{1-\beta} \mathcal{L}(t) \int_{t^{\gamma-1}}^1 \frac{dz}{z^\beta} =$$

$$= (1-\varepsilon)^2 \frac{t^{1-\beta}}{1-\beta} \mathcal{L}(t) \left(1 - t^{(\gamma-1)(1-\beta)} \right).$$

Similarly, we obtain the upper inequality. Thus, as $t \to \infty$

$$I_2(t) \sim \frac{t^{1-\beta}}{1-\beta} \mathcal{L}(t).$$

- For the first integral one can notice that $I_1(t) \leq t^\gamma$, because $\overline{B}(x) \leq 1$ for all x. Therefore, if we choose $0 < \gamma < 1 - \beta$ then the principal term of the asymptotics $I(t)$ is $t^{1-\beta}$.

Combining these estimates, we have

$$I(t) \sim \frac{t^{1-\beta}}{1-\beta} \mathcal{L}(t) \text{ as } t \to \infty.$$

Since $\mu_{ii}^T = \lambda I(T t_i)$, then for any $i \geq 1$

$$\mu_{ii}^T \sim \frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta} t_i^{1-\beta} \quad (5)$$

as $T \to \infty$.

The assertion of the *Lemma 2* follows from the representation for α_{ij}^T and given asymptotics (4) and (5). ■

Lemma 3 *Let the vector* $\boldsymbol{d} = (d_{11}, d_{12} \dots d_{1n}, d_{22}, d_{23} \dots d_{2n}, \dots, d_{nn})$. *Denote*

$$\widehat{\boldsymbol{\xi}}^T = \frac{\boldsymbol{\xi^T} - \frac{\lambda}{1-\beta} T^{1-\beta} \mathcal{L}(T) \boldsymbol{d}}{\sqrt{\frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta}}}.$$

If the Condition 2 is satisfied, then

$$\widehat{\boldsymbol{\xi}}^T \xrightarrow{d} \mathcal{N}(0, D) \text{ as } T \to \infty. \tag{6}$$

Here $\boldsymbol{\xi}^T = \left(\xi_{11}^T, \xi_{12}^T \dots \xi_{1n}^T, \xi_{22}^T, \xi_{23}^T \dots \xi_{2n}^T, \dots, \xi_{nn}^T\right)$, *$D$ is a diagonal matrix with following values on its main diagonal*

$$\{d_{11}, d_{12} \dots d_{1n}, d_{22}, d_{23} \dots d_{2n}, \dots, d_{nn}\}$$

where $\{d_{ij}\}_{i,j=1}^n$ *are defined in the Lemma 2.*

Proof. The assertion follows from the *Lemma 2* and the following property of a Poisson distribution. If a random variable ζ_λ has a Poisson distribution with parameter λ and $\zeta_\lambda^\beta = \beta \frac{\zeta_\lambda - \lambda}{\sqrt{\lambda}}$ then

$$\zeta_\lambda^\beta \xrightarrow{d} \mathcal{N}(0, \beta^2) \text{ as } \lambda \to \infty.$$

Since ξ_{ij}^T has a Poisson distribution with the parameter α_{ij}^T and $\alpha_{ij}^T \to \infty$ by the *Lemma 2* as $T \to \infty$ then

$$\frac{\xi_{ij}^T - \frac{\lambda}{1-\beta} T^{1-\beta} \mathcal{L}(T) d_{ij}}{\sqrt{\frac{\lambda}{1-\beta} T^{1-\beta} \mathcal{L}(T)}} \xrightarrow{d} \mathcal{N}(0, d_{ij}) \text{ as } T \to \infty$$

where $1 \le i \le n$, $1 \le j \le n$.

There is a convergence of vectors, since random variables $\left\{\xi_{ij}^T\right\}_{i,j=1}^n$ are independent. So, the *Lemma 3* is proved. ■

Now we are ready to formulate and prove the analogue of the *Theorem 1* for the system S_1.

Theorem 2 *If for the queuing system* S_1 *the Condition 2 is satisfied, then*

$$\frac{\boldsymbol{q}^T - \boldsymbol{\rho} \frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta}}{\sqrt{\frac{\lambda}{1-\beta} \mathcal{L}(T) T^{1-\beta}}} \xrightarrow{d} \mathcal{N}(0, R) \, as \, T \to \infty.$$

Here the vector $\boldsymbol{\rho} = (t_1^{1-\beta}, t_2^{1-\beta}, \dots, t_n^{1-\beta})$ *and the covariance matrix* $R = (r_{ij})$ *consists of elements* $r_{ij} = t_j^{1-\beta} - (t_j - t_i)^{1-\beta}$ *for* $1 \le i \le j \le n$.

Proof. One can notice that for any $1 \le k \le n$

$$q(Tt_k) = \sum_{i=1}^{k} \sum_{j=k}^{n} \xi_{ij}^T,$$

so that a vector $\boldsymbol{q}^T = (q(Tt_1), q(Tt_2), \dots, q(Tt_n))$ has the representation

$$\boldsymbol{q}^T = C\boldsymbol{\xi}^T.$$

Here C is the matrix with constant coefficients of size $\frac{n(n+1)}{2} \times n$. For example, for $n = 4$ matrix C has the form

$$\begin{pmatrix} 1\,1\,1\,1\,0\,0\,0\,0\,0\,0 \\ 0\,1\,1\,1\,1\,1\,1\,0\,0\,0 \\ 0\,0\,1\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,0\,1\,0\,0\,1\,0\,1\,1 \end{pmatrix}.$$

We note that a normalized vector

$$\widehat{\boldsymbol{q}^T} = \frac{\boldsymbol{q^T} - \frac{\lambda}{1-\beta} T^{1-\beta} \mathcal{L}(T) C \boldsymbol{d}}{\sqrt{\frac{\lambda}{1-\beta} T^{1-\beta} \mathcal{L}(T)}}$$

can be rewritten as $\widehat{\boldsymbol{q}^T} = C\widehat{\boldsymbol{\xi}^T}$. Since $\widehat{\boldsymbol{\xi}^T} \xrightarrow{d} \xi^\infty$ as $T \to \infty$, then $\widehat{\boldsymbol{q}^T} \xrightarrow{d} C\xi^\infty$ as $T \to \infty$. According to the *Lemma 3* random vector ξ^∞ is the normal distributed.

Here we take into account the following property of the Gaussian distribution. If a vector X has a normal distribution with parameters (μ, Σ), then for any matrix A the vector AX has a normal distribution with parameters $(A\mu, A\Sigma A^T)$. So, the vector $C\xi^\infty$ has a normal distribution with zero mean and some covariance matrix. To find the form of this matrix we note that for $i \le k$ we have the following equality

$$cov(q(Tt_i), q(Tt_k)) = \lambda \int_0^{Tt_i} \overline{B}(Tt_k - y) dy = \mu_{ik},$$

therefore hereinafter all follows from the proof of the *Lemma 2*.■

It is evident that the *Theorem 2* is a corollary of the *Theorem 1*.

4 Proof of the main result

Assume that the intensity of the input flow $\lambda(t)$ is a function which satisfies the *Condition 1*. We note that statements of lemmas which proved in the previous section, remain valid in this case. Let us consider them one after another. The assertion of the *Lemma 1* is retained with the only difference that in this case

$$\alpha_{ij}^T = \int_{Tt_{i-1}}^{Tt_i} \left(\overline{B}(Tt_j - y) - \overline{B}(Tt_{j+1} - y) \right) \lambda(y) dy.$$

The *Lemma 3* is formulated in the same way as in the case of a constant intensity. On the proof of the *Lemma 2* we will focus in more detail.

Lemma 4 *If Conditions 1 and 2 are fulfilled then for a function α_{ij}^T the statement of the Lemma 2 takes place.*

Proof. We denote

$$m_{ij}^T = \int_0^{Tt_i} \overline{B}(Tt_j - y)\lambda(y)dy.$$

Using this notation we can rewrite α_{ij}^T as

$$\alpha_{ij}^T = m_{ij}^T - m_{i-1j}^T - m_{ij+1}^T + m_{i-1j+1}^T. \tag{7}$$

Thus, to find the asymptotic behavior of α_{ij}^T as $T \to \infty$, it is sufficient to find the asymptotic behavior for m_{ij}^T as $T \to \infty$, $1 \le i \le j \le n$. We prove this lemma just like the *Lemma 2* in two steps.

1. Let $i < j$. In this case we only need to obtain asymptotics of the integral for $i = 1$, $j = 2$

$$J_T = \int_0^{Tt_1} \overline{B}(Tt_2 - y)\lambda(y)dy.$$

Let $N(T) = \max\{k : S_k < Tt_1\}$. In view of the monotonicity of the function $\overline{B}(Tt_2 - y)$ we have

$$J_T \le J_T^+ = \sum_{k=0}^{N(T)-1} \overline{B}(Tt_2 - S_{k+1}) \int_{S_k}^{S_{k+1}} \lambda(y)dy = \lambda \sum_{k=0}^{N(T)-1} \overline{B}(Tt_2 - S_{k+1})(S_{k+1} - S_k).$$

Since $t_2 > t_1$ then for any $\varepsilon > 0$ there is $T_\varepsilon^{(1)}$ such that

$$\overline{B}(Tt_2 - S_{k+1}) \le (1+\varepsilon)\frac{\mathcal{L}(Tt_2 - S_k)}{(Tt_2 - S_k)^\beta}$$

for all $k \ge 0$. Therefore

$$J_T^+ \le \lambda(1+\varepsilon)T \sum_{k=1}^{N(T)-1} \frac{\mathcal{L}(Tt_2 - S_k)}{T^\beta \left(t_2 - \frac{S_k}{T}\right)^\beta} \frac{S_{k+1} - S_k}{T} =$$

$$= \lambda(1+\varepsilon)T^{1-\beta}\mathcal{L}(T) \sum_{k=1}^{N(T)-1} \frac{\mathcal{L}(T(t_2 - \frac{S_k}{T}))}{\mathcal{L}(T)} \frac{1}{\left(t_2 - \frac{S_k}{T}\right)^\beta} \frac{S_{k+1} - S_k}{T}.$$

This implies that there is $T_\varepsilon^{(2)}$ such that for $T > \max(T_\varepsilon^{(1)}, T_\varepsilon^{(2)})$

$$J_T^+ \le \lambda(1+\varepsilon)^2 T^{1-\beta}\mathcal{L}(T) \sum_{k=0}^{N(T)-1} (t_2 - r_k)^\beta (r_{k+1} - r_k)$$

where $r_k = \frac{S_k}{T}$. Then

$$\sum_{k=0}^{N(T)} \frac{r_{k+1} - r_k}{(t_2 - r_k)^\beta} \to \int_0^{t_1} \frac{dy}{(t_2 - y)^\beta} = \frac{1}{1-\beta}\left[t_2^{1-\beta} - (t_2 - t_1)^{1-\beta}\right] \text{ as } T \to \infty.$$

we have

$$J_T^+ \leq \frac{\lambda}{1-\beta}(1+\varepsilon_1)T^{1-\beta}\mathcal{L}(T)$$

for any $\varepsilon_1 > 0$ and all sufficiently large T. Similarly the lower estimate is obtained. So, for any $i < j$ we get the following equivalence

$$m_{ij}^T \sim \frac{\lambda}{1-\beta}\left(t_j^{1-\beta} - (t_j - t_i)^{1-\beta}\right)T^{1-\beta}\mathcal{L}(T) \text{ as } T \to \infty.$$

2. Now we consider the case when $i = j$. Let us denote $I(t) = \int\limits_0^t \overline{B}(y)\lambda(t-y)dy$, then $\mu_{ii}^T = I(Tt_i)$. We find the asymptotic behavior of the function $I(t)$ as $t \to \infty$. Since for any $0 < \gamma < 1$ the integral $I(t)$ can be represented as follows

$$I(t) = \int_0^{t^\gamma} \overline{B}(y)\lambda(t-y)dy + \int_{t^\gamma}^t \overline{B}(y)\lambda(t-y)dy = I_1(t) + I_2(t).$$

Let us find the asymptotic behavior of each of the terms as $t \to \infty$.

- Arguing similarly to the first part of this lemma (the case of $i < j$), we get

$$I_2(t) \sim \frac{\lambda}{1-\beta}t^{1-\beta}\mathcal{L}(t) \text{ as } t \to \infty.$$

- Further, we note that $I_1(t) \leq \int\limits_0^{t^\gamma} \lambda(t-y)dy = \Lambda(t) - \Lambda(t - t^\gamma) \sim \lambda t^\gamma$ as $t \to \infty$.

Thus, if we choose $0 < \gamma < 1-\beta$, then the principal term of the asymptotic for $I(t)$ will be $t^{1-\beta}$. So, we obtain that for any $1 \leq i \leq n$

$$m_{ii}^T \sim \frac{\lambda}{1-\beta}t_i^{1-\beta}\mathcal{L}(T)T^{1-\beta}.$$

The *Lemma 4* implies from obtained asymptotics for m_{ij}^T and representation (7) for α_{ij}^T.■

Further using notes given in the case of constant intensity we obtain the *Theorem 1.*

Remark 1 *For the system S with a Poisson input flow with intensity $\lambda(t)$ where $\lambda(t)$ is a periodic function with a period $\tau > 0$. If the Condition 2 is satisfied, then the Theorem 1 holds and $\lambda = \frac{\Lambda(\tau)}{\tau}$.*

Author is deeply grateful to Professor L. G. Afanas'eva for her constant interest in this work, valuable suggestions and guidance, significantly contributed to the writing this article.

References

1. L.G. Afanasyeva, A.P. Bilenko, *Characteristic of total flow of impulses*, in Questions of electronics, Issue 4, 1972 (in Russian).
2. L.G. Afanasyeva, I.V. Rudenko , *Queuing systems $GI|G|\infty$ and their applications to the analysis of traffic patterns*, Theory of Probability and its Applications, 57, 3: 427–452, 2012.
3. A.D. Ventcel, *Course of the Theory of the Stochastic Processes.* Nauka, Moscow, 1975 (in Russian)
4. A.A. Borovkov *Asymptotic methods in queuing theory.* John Wiley Sons, 1, 1984.
5. E. Seneta *Regularly varying functions in the theory of simple branching processes* Advances in Applied Probability. 408–420, 1974.
6. DAuria, Bernardo. *Stochastic decomposition of the queue in a random environment.* Operations Research Letters 35, 6: 805-12, 2007.
7. Kang, Weining, Kavita Ramanan,*Asymptotic Approximations for Stationary Distributions of Many-Server Queues with Abandonment.* The Annals of Applied Probability 22, 2, 477-521, 2012.
8. N. Kaplan *Limit theorems for a $GI/G/\infty$ queue.* The Annals of Probability, Vol.3,No.5, 780–789, 1975.
9. L. Lipsky, D. Derek, G. Swapna, *New frontiers in applied probability*, 2011.
10. Massey, William A., Ward Whitt. *Networks of Infinite-Server Queues with Nonstationary Poisson Input.* Queueing Systems 13, 13: 183-250, 1993.
11. W. Whitt *Efficiency-Driven Heavy-Traffic Approximations for Many-Server Queues with Abandonments.* Management Science 50, 10: 1449-61, 2004.
12. M. Ya. Postan, *Flow of Serviced Requests in Infinite-Channel Queueing Systems in a Transient Mode*, Probl. Peredachi Inf., 13:4, 8995, 1977.

Claim reserving including risk margins

Teresa Costa[1] and Eva Boj[2]

[1] Departamento de Matemática Económica, Financiera y Actuarial. Universidad de Barcelona
(E-mail: tcosta@ub.edu)

[2] Departamento de Matemática Económica, Financiera y Actuarial. Universidad de Barcelona
(E-mail: evaboj@ub.edu)

Abstract. We study different ways to include risk margins to the estimation of provisions in the IBNR claim reserving problem with run-off triangles. There are different types of amounts which are of interest: the provisions for the different accident years, the future payments for the different calendar years and the total provision. We are interested in to calculate the present value of the future payments for the different calendar years, taking into account the Solvency II Directive. In this context, firstly we can calculate the present value using a risk-free interest rate term structure. Secondly, we can work in a conservative scenario by adding risk margins to the fitted payments by calendar years. Assume that we are using a stochastic claim reserving method as the generalized linear model (GLM) and that we are able to estimate prediction errors and predictive distributions by bootstrapping. There are different ways to include risk margins: one could be to add a percentage of the prediction errors to the fitted values, or another one could be to calculate the value-at-risk with a given confidence level of the predictive distributions of the future payments by calendar years. Finally, we can compute its present values in all cases.

Keywords: Technical provisions, Generalized Linear Model, Solvency II, Calendar years.

1 Introduction

Assume a GLM (see, e.g., Boj *et al*. [2]; Boj and Costa [3] and McCullagh and Nelder [12]) to model the incremental losses of a run-off triangle, c_{ij} . Assume a parametric family for the error distribution with mean and variance:

$$\mu_{ij} = E\left[c_{ij}\right] \quad \text{and} \quad Var\left[c_{ij}\right] = \left(\phi/w_{ij}\right)V\left(\mu_{ij}\right) = \left(\phi/w_{ij}\right)\mu_{ij}^{\theta}, \quad (1)$$

The variance function depends on a parameter θ and this parametric family has as particular cases: the Poisson distribution when $\theta = 1$; the Gamma distribution when $\theta = 2$; and the Inverse Gaussian distribution when $\theta = 3$.
Assuming the logarithmic link function

$$\log \mu_{ij} = \eta_{ij}, \quad (2)$$

16th ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece

the predicted values $\hat{c}_{ij}$ are estimated from $\hat{c}_{ij} = \exp\left(\hat{c}_0 + \hat{\alpha}_i + \hat{\beta}_j\right)$, where α_i is the factor corresponding to the accident year $i = 1,...,k$ and β_j is the factor corresponding to the development year $j = 1,...,k$. The c_0 value is the term corresponding to the accident year 0 and development year 0.

The future payments for the different calendar years $t = k+1,...,2k$ are obtained by adding the incremental losses that were made in the future calendar years:

$$FP_t = \sum_{j=t-k}^{k} \hat{c}_{t-j,j} \, . \quad (3)$$

We calculate the present value of the future payments by calendar years, as it is indicated in the Solvency II Directive (see Albarrán and Alonso [1]). Additionally, we describe two ways of consider a risk margin to obtain the IBNR claim provisions. We use the prediction errors for the future payments by calendar years and also consider the estimated value at risk (VaR) of the predictive distribution for the future payments by calendar years at a given significance level.

2 Prediction errors for future payments by calendar years

The prediction error (PE) of the parametric family of distributions (1) assuming (2) has been studied in the cases of the accident years provisions and the total provision in, e.g., England and Verrall [7], [8], [10] and England [9]. The study is extended to the IBNR future payments by calendar years in Boj *et al*. [4] and Espejo *et al*. [11] for the particular case in which the over-dispersed Poisson distribution is assumed.

Now we extend the formulas of the prediction error to the future payments by calendar years for the general parametric family (1) assuming (2).

The PE for the future payments by calendar years with the analytical formula of the distribution can be calculated as the squared root of the Mean Squared Error (MSE) of prediction:

$$MSE\left(FP_t\right) = E\left[\left(FP_t - FP_t\right)^2\right] \approx \sum_{\substack{i,j=1,..,k \\ i+j=t}} \phi\mu_{ij}^{\theta} + \mu_t^T Var\left[\eta_t\right]\mu_t =$$

$$= \sum_{\substack{i,j=1,..,k \\ i+j=t}} \phi\mu_{ij}^{\theta} + \sum_{\substack{i,j=1,..,k \\ i+j=t}} \mu_{ij}^{2} Var\left[\eta_{ij}\right] + 2 \sum_{\substack{i_1,i_2,j_1,j_2=1,..,k \\ i_1 j_1 \neq i_2 j_2 \\ i_1+j_1=t, i_2+j_2=t}} \mu_{i_1 j_1}\mu_{i_2 j_2} Cov\left[\eta_{i_1 j_1}, \eta_{i_2 j_2}\right]. \quad (4)$$

$$t = k+1,...,2k$$

Alternatively, we can estimate the predictive distribution of c_{ij} by bootstrap methodology (see Efron and Tibshirani [5]); in the case of the GLM and the

IBNR claim provisions it is usual to apply bootstrapping residuals based on Pearson residuals.

In the bootstrap process estimation we have *B* resamples and, from these values, we can calculate the squared SE, i.e. the variance of the predictive distribution for the future payments by calendar years.
The bootstrap estimations of the PE for the future payments by calendar years are:

$$PE^{boot}(FP_t) \approx \sqrt{\sum_{\substack{i,j=1,..,k \\ i+j=t}} \hat{\phi}^P \hat{c}_{ij}^{\theta} + SE\left(FP_t^{boot}\right)^2}\ , \quad t = k+1,...,2k \ . \quad (5)$$

3 Application

To illustrate the proposed methodology we use the triangle of Taylor and Ashe [15] in Figure 1, with n=55 incremental losses for accident years $i = 0,\ldots,9$ and development years $j = 0,\ldots,9$.

i	*j*									
	0	1	2	3	4	5	6	7	8	9
0	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
1	352118	884021	933894	1183289	445745	320996	527804	266172	425046	
2	290507	1001799	926219	1016654	750816	146923	495992	280405		
3	310608	1108250	776189	1562400	272482	352053	206286			
4	443160	693190	991983	769488	504851	470639				
5	396132	937085	847498	805037	705960					
6	440832	847361	1131398	1063269						
7	359480	1061648	1443370							
8	376686	986608								
9	344014									

Figure 1. Run-off triangle of Taylor and Ashe [15] with 55 incremental losses

This dataset has been used in many texts on IBNR problems (see Boj and Costa [5], England and Verrall [7], England [9] and Renshaw [13], [14]. In England and Verrall [7] the coefficient of variation (the PE as a percentage of the provision estimate) is calculated for the accident years and total provisions in the following cases: when the over-dispersed Poisson and the Gamma distributions are assumed in the analytic formulas to calculate PE and when the over-dispersed Poisson distribution is assumed in the bootstrap methodology for the estimation of PE.

We include in this section the computation of the present values of the future payments by calendar years adding a risk margin in two ways. First, we calculate the present value of the future payments by calendar year plus a percentage δ of the PE, with analytic formula and with bootstrap estimation. We consider a percentage of $\delta = 0.25$ as in Boj and Costa [5]. Second, we calculate the present value of the values at risk of the predictive distribution for the future payments by calendar year at a level of confidence equal to 99.5%, following the recommendations of Solvency II.

We assume the over-dispersed Poisson and the Gamma distributions with the logarithmic link function in the GLM and in the bootstrapping estimations we obtain B=1000 resamples. In the nine future calendar years it is assumed a fixed annual interest rate equal to 1.5%. We have used the R software for the computations and, specifically, the *glm* function of the *stats* package.

The numerical results are shown in Tables 1 to 8. If we compare the coefficients of variation, the values in the case of the Gamma distribution (when $\theta = 2$) are lower than in the case of the over-dispersed Poisson distribution (when $\theta = 1$).

Calendar year	Payment	Prediction error	Coefficient of variation
10	5226535.8	747369.6	14.30 %
11	4179394.4	710144.6	16.99 %
12	3131667.5	644139.5	20.57 %
13	2127271.9	479125.6	22.52 %
14	1561878.9	404967.7	25.93 %
15	1177743.7	364294.9	30.93 %
16	744287.4	294424.6	39.56 %
17	445521.3	250986.8	56.34 %
18	86554.6	108268.8	125.09 %

Table 1. Future payments by calendar years, prediction errors and coefficients of variation for the over-dispersed Poisson distribution using analytic formula.

Calendar year	Payment	Prediction error	Coefficient of variation
10	5096855.3	847281.6	16.62 %
11	4050001.5	749549.8	18.51 %
12	3064407.7	628141.0	20.50 %
13	2078010.5	431885.8	20.78 %
14	1510392.7	345880.7	22.90 %
15	1095402.7	292255.7	26.68 %
16	692118.4	220057.8	31.79 %
17	416539.9	181226.5	43.51 %
18	82075.9	47918.1	58.38 %

Table 2. Future payments by calendar years, prediction errors and coefficients of variation for the Gamma distribution using analytic formula.

Calendar year	Mean payment	Standard deviation	Prediction error	Coefficient of variation

10	5262187.6	748660.0	756563.2	14.48 %
11	4206004.9	718618.2	721067.3	17.25 %
12	3153556.8	653134.1	649753.2	20.75 %
13	2139244.8	504395.7	487995.9	22.94 %
14	1562523.4	408073.2	411005.7	26.31 %
15	1178586.1	364102.8	365547.7	31.04 %
16	771451.6	302919.9	292974.4	39.36 %
17	455633.0	250906.3	254458.2	57.11 %
18	91579.0	104329.2	107988.8	124.76 %

Table 3. Future payments by calendar years, standard deviations, prediction errors and coefficients of variation for the over-dispersed Poisson distribution using bootstrap methodology with 1000 resamples.

Calendar year	Mean payment	Standard deviation	Prediction error	Coefficient of variation
10	5096897.2	1017.5	652964.8	12.81 %
11	4050047.9	1014.2	545647.9	13.47 %
12	3064465.1	920.9	434825.7	14.19 %
13	2078021.4	694.1	297581.3	14.32 %
14	1510393.3	580.7	233914.7	15.49 %
15	1095419.1	493.1	194252.2	17.73 %
16	692130.4	395.9	142601.5	20.60 %
17	416548.9	342.7	109441.6	26.27 %
18	82080.9	152.7	26649.2	32.47 %

Table 4. Future payments by calendar years, standard deviations, prediction errors and coefficients of variation for the Gamma distribution using bootstrap methodology with 1000 resamples.

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction error
10	1	5226535.8	5413378.2
11	2	4179394.4	4356930.6
12	3	3131667.5	3292702.4
13	4	2127271.9	2247053.3
14	5	1561878.9	1663120.8
15	6	1177743.7	1268817.4
16	7	744287.4	817893.5
17	8	445521.3	508268.0
18	9	86554.6	113621.8
Present value		17873967	18820197

Table 5. Present values of the future payments by calendar years and of the future payments by calendar years plus a 25% of the prediction error for the over-dispersed Poisson distribution using analytic formula and assuming a 1.5% fixed annual interest rate.

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction error
10	1	5096855.3	5308675.7
11	2	4050001.5	4237389.0
12	3	3064407.7	3221442.9
13	4	2078010.5	2185982.0
14	5	1510392.7	1596862.9
15	6	1095402.7	1168466.7
16	7	692118.4	747132.9
17	8	416539.9	461846.5
18	9	82075.9	94055.4
Present value		17310125	18199962

Table 6. Present values of the future payments by calendar years and of the future payments by calendar years plus a 25% of the prediction error for the Gamma distribution using analytic formula and assuming a 1.5% fixed annual interest rate.

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction Error	$VaR_{99.5}$
10	1	5226535.8	5415676.6	7417055.0
11	2	4179394.4	4359661.3	6364764.7
12	3	3131667.5	3294105.8	5207534.8
13	4	2127271.9	2249270.9	3682095.3
14	5	1561878.9	1664630.3	2735533.8
15	6	1177743.7	1269130.6	2209257.2
16	7	744287.4	817531.0	1841310.7
17	8	445521.3	509135.8	1262432.7
18	9	86554.6	113551.8	473412.3
Present value		17873967	18830614	29688278

Table 7. Present values of the future payments by calendar years, of the future payments by calendar years plus a 25% of the prediction error and of the Value at Risk with a confidence level of the 99.5% for the over-dispersed Poisson

distribution using bootstrap methodology with 1000 resamples and assuming a 1.5% fixed annual interest rate.

Calendar year	Deferral (in years)	Payment	Payment + 0.25 Prediction Error	$VaR_{99.5}$
10	1	5096855.3	5260096.5	5099652.6
11	2	4050001.5	4186413.5	4052656.1
12	3	3064407.7	3173114.1	3067038.9
13	4	2078010.5	2152405.9	2079754.3
14	5	1510392.7	1568871.3	1511867.6
15	6	1095402.7	1143965.8	1096662.3
16	7	692118.4	727768.8	693145.23
17	8	416539.9	443900.3	417478.6
18	9	82075.9	88738.2	82457.2
Present Value		17310125	17938348	17324230

Table 8. Present values of the future payments by calendar years, of the future payments by calendar years plus a 25% of the prediction error and of the Value at Risk with a confidence level of the 99.5% for the Gamma distribution using bootstrap methodology with 1000 resamples and assuming a 1.5% fixed annual interest rate.

Conclusions

We have extended the formulas of the prediction errors for the accident years provisions and total provision in England and Verrall [7], [8], [10] and England [9] to the future payments by calendar years in the case of the general parametric family of error distributions with logarithmic link.
The present value of the future payments by calendar years allows at the actuary calculate the best estimate of the technical provisions in the context of Solvency II. We include risk margins by means of the prediction errors and the VaR of the predictive distribution for the future payments by calendar years.

Acknowledgements

Authors have been supported by the Spanish Ministerio de Educación y Ciencia under grant MTM2014-56535-R and by the Generalitat de Catalunya, AGAUR under grant 2014SGR152.

References

1. I. Albarrán and P. Alonso. *Métodos estocásticos de estimaciones de las provisiones técnicas en el marco de Solvencia II*. Cuadernos de la Fundación MAPFRE, 158. Fundación MAPFRE Estudios, Madrid, 2010.
2. E. Boj, M.M. Claramunt and J. Fortiana. *Análisis multivariante aplicado a la selección de factores de riesgo en la tarificación*. Cuadernos de la Fundación MAPFRE, 88. Fundación MAPFRE Estudios, Madrid, 2004.
3. E. Boj and T. Costa. *Modelo lineal generalizado y cálculo de la provisión técnica*. Depósito digital de la Universidad de Barcelona. Colección de objetos y materiales docentes (OMADO), 2014. http://hdl.handle.net/2445/49068
4. E. Boj, T. Costa and J. Espejo. Provisiones técnicas por años de calendario mediante modelo lineal generalizado. Una aplicación con RExcel. *Anales del Instituto de Actuarios Españoles*, Tercera Época, 20, 83–116, 2014.
5. E. Boj and T. Costa. Claim reserving: calendar year reserves for the GLM. In: *Current Topics on Risk Analysis: ICRA6 and RISK2015 Conference*. Cuadernos de la Fundación MAPFRE , 205, pp. 169–177. Fundación MAPFRE, Madrid, 2015.
6. B. Efron. and J. Tibshirani. *An Introduction to the bootstrap*. New York: Chapman & Hall/CRC, New York, 1998.
7. P. D. England and R.J. Verrall. Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25, 281–293, 1999.
8. P.D. England and R.J. Verrall. Stochastic claims reserving in general insurance (with discussion). *British Actuarial Journal*, 8, 443–544, 2002.
9. P.D. England. Addendum to “Analytic and bootstrap estimates of prediction errors in claim reserving”. *Insurance: Mathematics and Economics*, 31, 461–466, 2002.
10. P.D. England and R.J. Verrall. Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1 (II), 221–270, 2006.
11. J. Espejo, E. Boj and T. Costa. *Una aplicación de RExcel para el cálculo de provisiones técnicas con modelo lineal generalizado*. Depósito digital de la Universidad de Barcelona. Colección de Investigación-Software, 2014. http://hdl.handle.net/2445/56230
12. P. McCullagh and J. Nelder. *Generalized linear models. Second edition*. Chapman and Hall, Londres, 1989.
13. A.E. Renshaw. Chain ladder and interactive modelling (claims reserving and GLIM). *Journal of the Institute of Actuaries*, 116 (III), 559–587, 1989.
14. A.E. Renshaw. On the second moment properties and the implementation of certain GLIM based stochastic claims reserving models. *Actuarial Research Paper*, 65. Department of Actuarial Science and Statistics, City University, London, 1994.
15. G. Taylor, G. and F.R. Ashe, F.R. Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics*, 23, 37–61, 1983.
16. J. van Eeghen. *Loss reserving methods*. Surveys of Actuarial Studies, 1. National Nederlanden, 1981.
17. R. Verrall. An investigation into stochastic claims reserving models and the Chain-ladder technique. *Insurance: Mathematics and Economics*, 26, 91–99, 2000.
18. R. Verrall and P. England. Comments on: “A comparison of stochastic models that reproduce Chain Ladder reserve estimates, by Mack and Venter”. *Insurance: Mathematics and Economics*, 26, 109–111, 2000.

The Study of Population Ageing in the Context of Adult Health Status in the Czech Republic

Kornélia Cséfalvaiová[1], Jana Langhamrová[2], and Jitka Langhamrová[3]

[1]Department of Demography, University of Economics, Prague, Czech Republic (E-mail: kornelia.csefalvaiova@vse.cz)
[2]Department of Statistics and Probability, University of Economics, Prague, Czech Republic (E-mail: jana.langhamrova@vse.cz)
[3]Department of Demography, University of Economics, Prague, Czech Republic (E-mail: jitka.langhamrova@vse.cz)

Abstract. As people live on average longer, and as life expectancy increases, we may assume that we live healthier for a longer period. In studies of population ageing not only the individual age itself is important, but also the number of years spent in good health. Similarly, when we examine the expected number of years we have before us, we are not interested only in the length of our remaining life, but also whether those years of life will be spent in a good health condition or with health limitations. The priority of each state should be the health status of its residents. For example, companies are well aware of the importance of the health status of their employees in terms of increasing productivity. Submitted paper focuses on the process of population ageing and adult health status in the Czech Republic in relation to mental diseases. Solving the health status of the elderly and finding the appropriate treatment for diseases in older age groups is one of the key objectives.

Keywords: Ageing, Health, Alzheimer's Disease, Dementia, Czech Republic.

1 Introduction

Population ageing belongs to the most discussed topics worldwide. What is more important, we need reliable and valid measures to describe population ageing and health-related outcomes in the population (COURAGE, 2012). Developed regions, including Europe, are facing the increasing proportion of people older than 60 years. What is more, population ageing is affecting developing countries as well, because of their decreasing fertility rates. We expect an increase in the proportion of people older than 65 years to 30% in 2060 (Börsch-Supan, 2013). The changes in the age structure of the population will lead to an increasing attention and importance of care need. Population ageing, longevity and reaching old ages lead to a high percentage of old people who are in a need of care. In last centuries, the proportion of people who survived to old age, was not so significant as in recent period.

16th ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece

2 Health and Life Expectancy

There are many indicators affecting health. Retirement, mortality, health care are strongly associated with the health status. When examining the quality of life, some information is provided by diseases and disabilities. Many old people above age 65 suffer from mental health problems. As the process of population ageing is spreading all around the world, the issue of quality of life in older ages is becoming highly important not only for the seniors themselves, but for public sector, health systems and national policies as well (Börsch-Supan, 2008). It is not inconsiderable whether added years to life are spent in a good health or with health limitations. Old age is becoming a universally achievable stage of life irrespective of structural factors such as class, gender or ethnicity (Petrová Kafková, 2013). On the one hand, we live longer and the life expectancy is prolonging. On the other hand, don't added years of life mean the increasing number of people suffering from illnesses occurring at old ages?

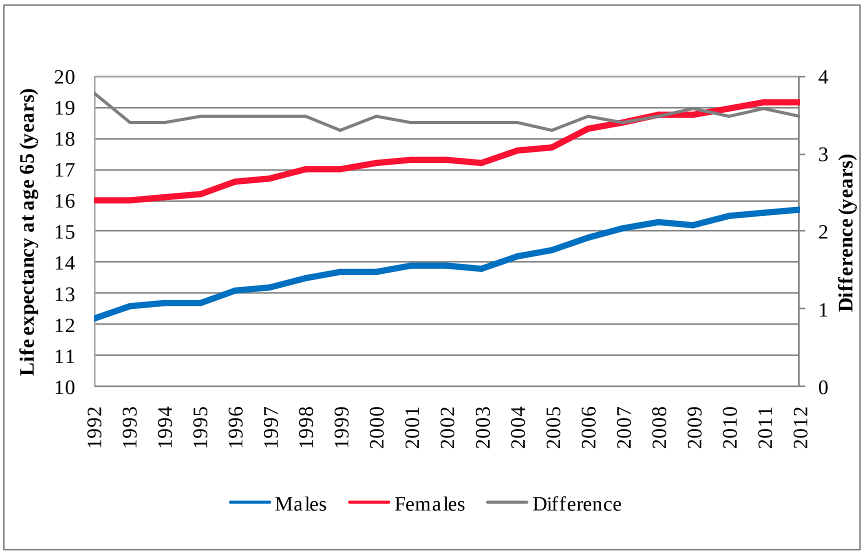

Fig. 1. Life expectancy at age 65 in the Czech Republic in the years 1992-2012
Source: Eurostat, authors' construction

In the last twenty years life expectancy at age 65 is increasing in the Czech Republic continuously. Difference between females and males was in average 3.5 years during this period (Fig. 1). The life expectancy at the age 65 was 12.2 years for males and 16 years for females in 1992. In 2012, the life expectancy at the age 65 was 15.7 years for males and 19.2 years for females (Fig. 1). Life expectancy and modal age at death belong among the most important demographic indicators from the view of adult mortality and longevity (Langhamrová et al., 2014). Many authors have been examining the area of life expectancy, see e.g. (Fiala, Langhamrová, 2014).

3 Alzheimer's Disease and Life Expectancy

Nowadays, mental, behavioral and emotional disorders belong to leading disabilities among older population aged 65 years and over. Due to higher life expectancy of women, women are more affected by dementia, Alzheimer's disease, Parkinson's disease and other neurological disorders (Fig. 2). Mortality statistics by causes of death was from 2007 affected by changes in the system of coding practices and updates classification, which is one of the innovations of the 10th revision. Objective was to improve the coding procedures in the process of selecting the underlying cause of death. These changes have affected mortality statistics, e.g. deaths by hypertension, heart failure, cerebrovascular diseases, atherosclerosis or diabetes (CZSO, 2014). From this reason we can see some changes in data in the period 2007–2010 (Fig. 2). We used data for the Czech Republic from the International classification of diseases (10th revision). As life expectancy at birth is prolonging, the number of males and females with mental disorders is also increasing. Life expectancy in 2012 for males was 75.1 years and for females 81.2 years. In comparison, number of deaths by Alzheimer’s disease was 774 females and 446 males. In 2003, it was only 459 females and 209 males. The number of deaths by Alzheimer's disease has almost doubled over the 10-year period. Which age groups contributed the most in the increase of deaths by Alzheimer’s disease? From Fig. 3 it is visible that it was for age groups 75-79, 80-84 and 85-89 for males and females. In younger age groups the occurrence of Alzheimer's disease is exceptional, it begins to detect after the age of 50 (Fig. 3).

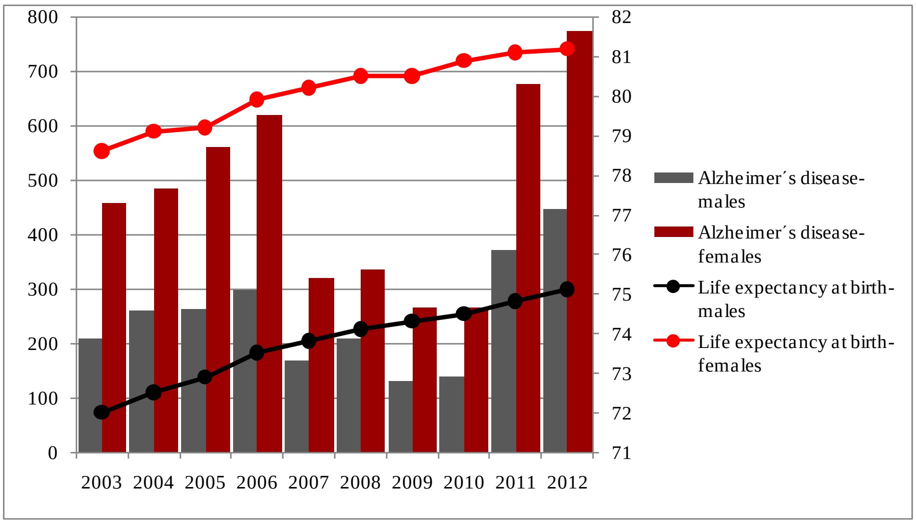

Fig. 2. Deaths by Alzheimer's disease and life expectancy at birth in the Czech Republic in the years 2003-2012

Source: CZSO, authors' construction

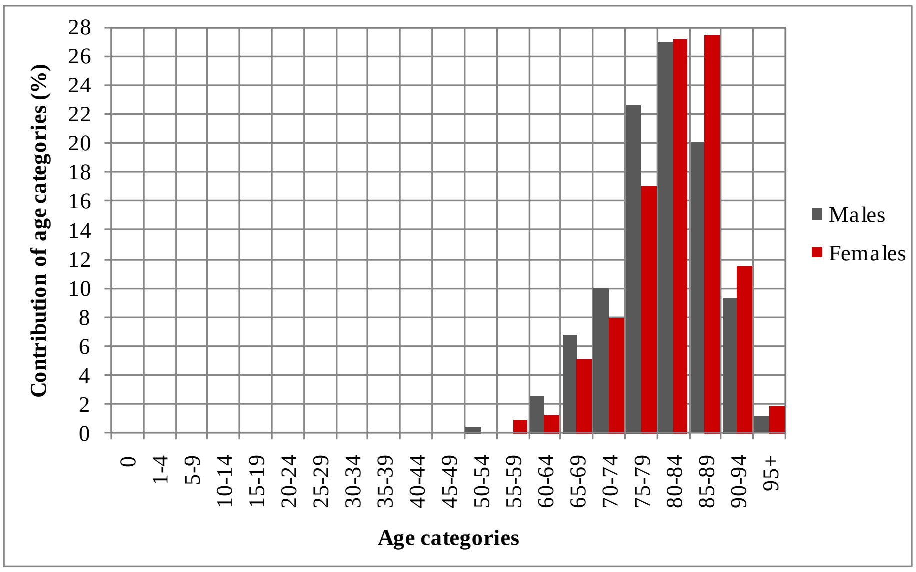

Fig. 3. Contribution of age categories to number of deaths by Alzheimer's disease in the Czech Republic in 2013
Source: CZSO, authors' construction

According to estimations, there are globally 35.6 million people with dementia with 7.7 million new cases every year–Alzheimer's disease is the most common form of dementia among older people and may contribute to 60–70% of cases (WHO, 2015). In the context of the increasing number of mentally ill patients, the number of mental health facilities will also increase. In the Czech Republic during the next two years there should be approximately 30 mental health centers established (Deník.cz, 2015).
What is more, an increased number of social workers will be needed, such as psychologists, psychiatrists, doctors, and other employees taking care of mentally incompetent patients.

4 Deaths by Nervous System Diseases

Nowadays, it is evident and without any discussions, that the length of people’s life is prolonging, mainly in the developed societies. We talk about decreasing death rates among old people. On the other hand, the more year people live, the more diseases can appear in their later stage of life. The importance of presence of nervous system diseases among old people is increasing. In the Czech Republic in the year 2003, there were altogether 2057 deaths by nervous system diseases, including 972 men and 1085 women. In 2013, there were altogether 2601 deaths by nervous system diseases, including 1163 men and 1438 women.

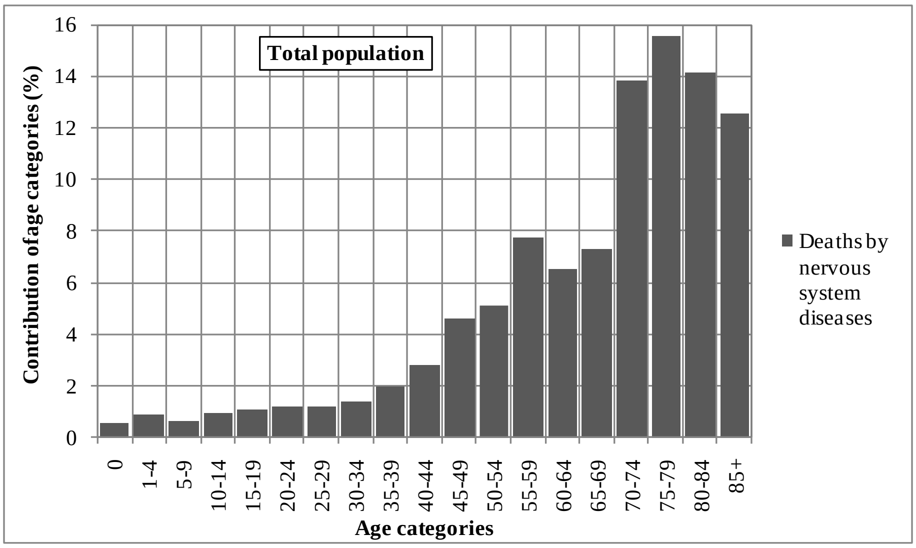

Fig. 4. Contribution of age categories to number of deaths by nervous system diseases in the Czech Republic in 2003

Source: CZSO, authors' construction

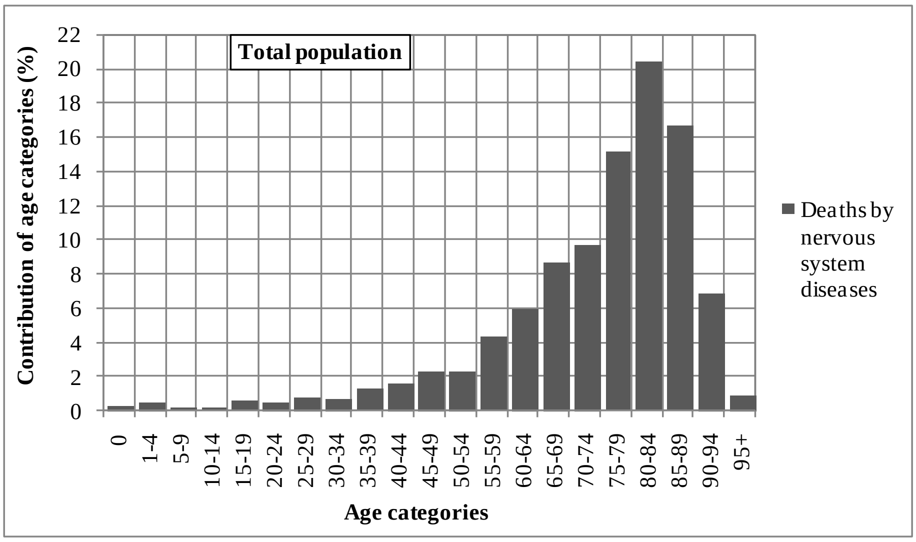

Fig. 5. Contribution of age categories to number of deaths by nervous system diseases in the Czech Republic in 2013

Source: CZSO, authors' construction

The difference makes 191 men and 353 women between 2003 and 2013 in the Czech Republic (Fig. 4 and Fig. 5). From these results it is evident that nervous system diseases are higher for women, but it is probably due to higher life expectancy of women. Probability of incidence of nervous system diseases is higher for women than for men because of their added years of life. From Fig. 4 it is visible, that in 2003 the number of deaths by nervous system diseases was the highest in age intervals 70-84. The contribution of the most important age category 75-79 was 15.5% in 2003. From Fig. 5 it is evident, that in 2013 the number of deaths by nervous system diseases has shifted to age categories 75-89. The contribution of the most important age category 80-84 was almost 20.5% in 2013. The difference between the year 2003 and 2013 makes 5 years and five-year age range.

5 Projections of the Future Prevalence of Dementia

For our projections of the prevalence of dementia we used the probabilities from EuroCode 2009 and EURODEM 1991. We supposed the same probabilities throughout the period and that the prevalence of dementia will be at the level of the year 2009 or 1991. We applied these probabilities on the projected age structure of the Czech population (we used the medium variant of the age structure projection from the Czech Statistical Office, 2013). We calculated the projections of the prevalence of dementia in the Czech Republic in the period 2013-2101 for men and women.

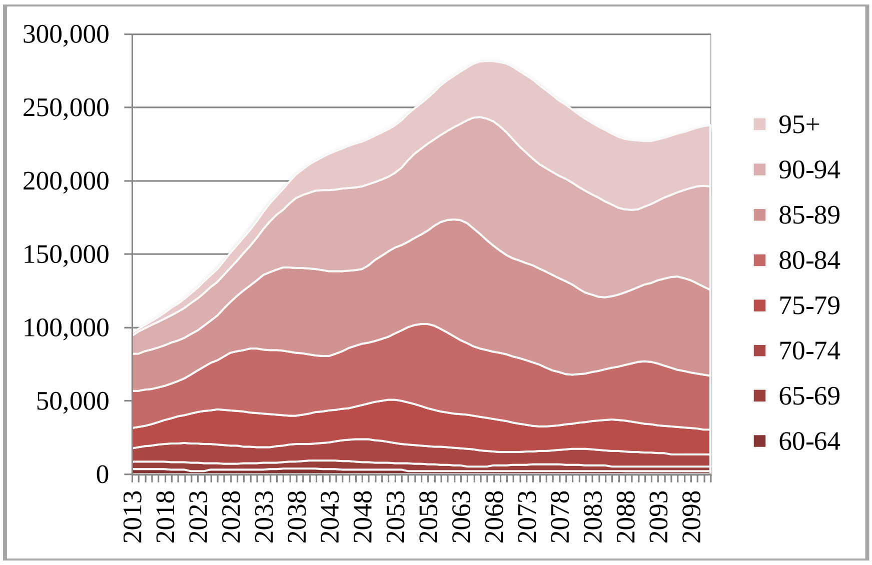

Fig. 6. Projection of the prevalence of dementia according to EuroCode 2009 (women)

Source: CZSO, authors' calculation

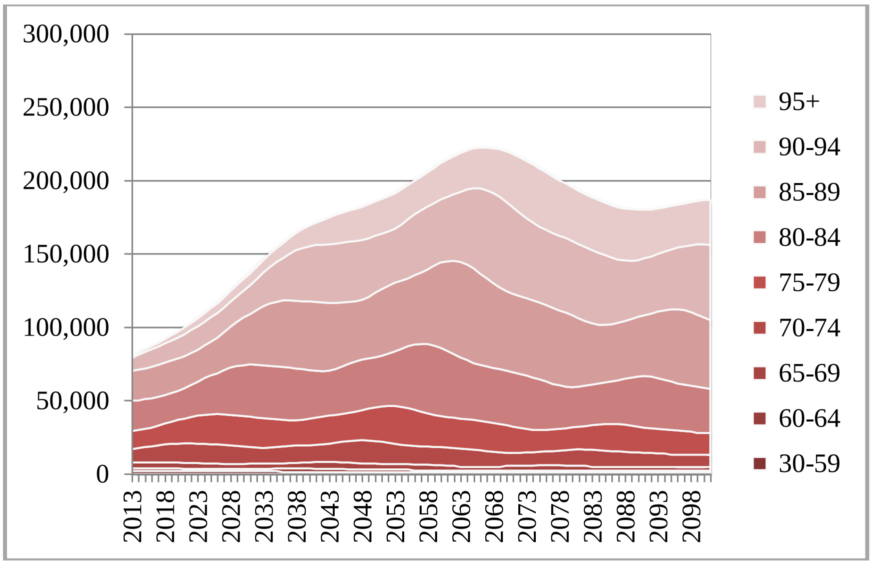

Fig. 7. Projection of the prevalence of dementia according to EURODEM 1991 (women)

Source: CZSO, authors' calculation

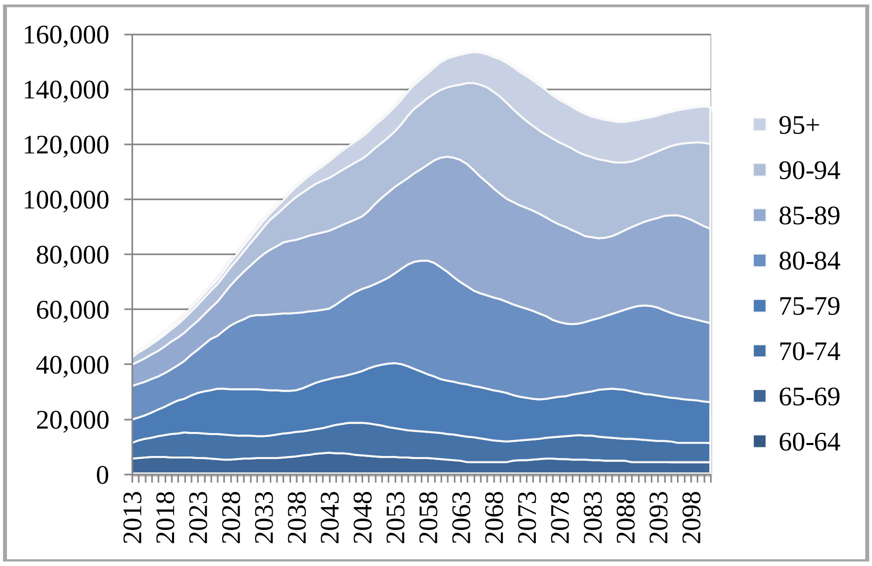

Fig. 8. Projection of the prevalence of dementia according to EuroCode 2009 (men)

Source: CZSO, authors' calculation

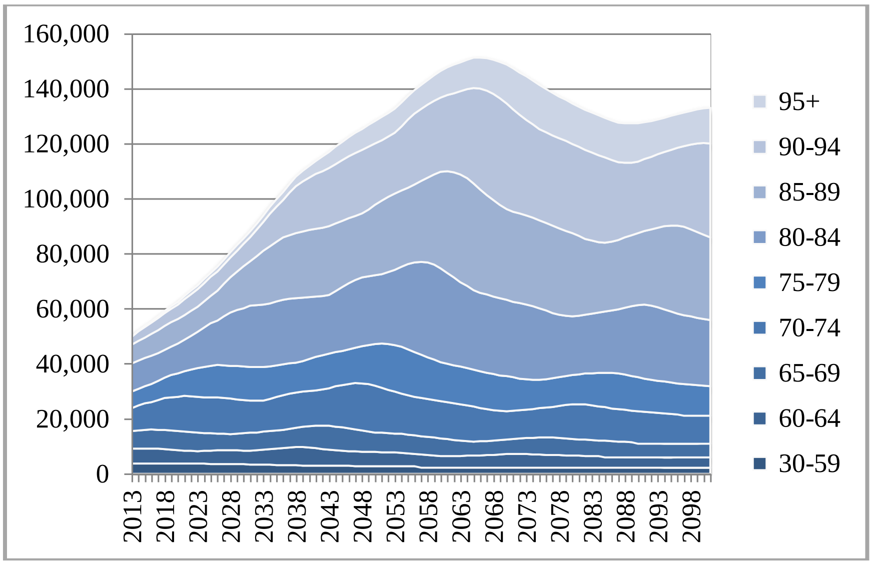

Fig. 9. Projection of the prevalence of dementia according to EURODEM 1991 (men)
Source: CZSO, authors' calculation

By comparing results for women, there will be almost 240 000 women suffering from dementia in 2101 (using probabilities from EuroCode 2009) and almost 190 000 women according to EURODEM 1991 (Fig. 6 and Fig. 7).
In case of men, the difference between using the probabilities from EuroCode 2009 and EURODEM 1991 is not so significant. According to both versions of projections, there will be almost 140 000 men suffering from dementia in the Czech Republic in 2101 (Fig. 8 and Fig. 9). When focusing on age categories that contribute the most to the prevalence of dementia, in 2013 the most important age group was 85-89 for women (according to EuroCode 2009) and 80-84 (according to EURODEM 1991). For men, it was the age interval 80-84 in both cases. In the future, the most significant age group will be 90-94 for women and 85-89 for men. A significant shift in the contribution of age groups is visible.

Conclusions

In our paper we focused on the process of population ageing in the Czech Republic from the view of deaths by nervous system diseases and the development of Alzheimer's disease among old population. Mental diseases, like Alzheimer's disease or Parkinson's disease, are significant mainly at older ages, after the age of 50. According to projections of the future prevalence of dementia, the share of men and women suffering from dementia will be increasing.

Population ageing undoubtedly represents an objective advance in the average length of human life. On the other hand, in the future, we can't be satisfied only with this occurred reality, but we have to focus on possible problems and their effective solutions. From this reason 21st century is the century of social services, social institutions, healthcare systems. Question of population ageing is affecting public institutions, but also family relatives, who will have to take care of the ageing family members.

Acknowledgement

This article was supported by the Internal Grant Agency of the University of Economics, Prague No. 68/2014 under the title "Economic and health connections of population ageing."

References

1. Börsch-Supan, A., A. Brugiavini, H. Jürges, A. Kapteyn, J. Mackenbach, J. Siegrist and G. Weber. (2008). First results from the Survey of Health, Ageing and Retirement in Europe (2004-2007).Starting the longitudinal dimension. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).
2. Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology DOI: 10.1093/ije/dyt088.
3. COURAGE. Collaborative Research on Ageing in Europe. 2012. http://www.courageineurope.eu/?page_id=28
4. CZSO. Czech Statistical Office. Deaths: by detailed list of causes of deaths, sex and age.http://www.czso.cz/csu/2014edicniplan.nsf/p/130067-14
5. Deník.cz. 2015. Do dvou let má v Česku vzniknout 30 centerduševníhozdraví. http://www.denik.cz/z_domova/do-dvou-let-ma-v-cr-vzniknout-30-center-dusevniho-zdravi-20150215.html
6. EuroCode. 2006-2008: European Collaboration on Dementia (EuroCoDe). 2012. http://www.alzheimer-europe.org/Alzheimer-Europe/Our-work/Completed-projects/2006-2008-EuroCoDe
7. EUROSTAT. European Statistical Office. 2015. http://ec.europa.eu/eurostat/data/database
8. Fiala, T., Langhamrová, J. Increase of Labor Force of Older Age - Challenge for the Czech Republic in Next Decade. In TSOUNIS, N. -- VLAHVEI, A. (ed.). International Conference on Applied Economics, ICOAE 2014. Amsterdam: Elsevier Ltd, 2014, s. 144-153. URL: http://www.sciencedirect.com/science/article/pii/S221256711400330X
9. Langhamrová, J., Cséfalvaiová, K., Langhamrová, J. Life Expectancy and Modal Age at Death in Selected European Countries in the Years 1950-2012. In *SMTDA 2014 Stochastic Modeling Techniques and Data Analysis International Conference and Demographics Workshop*. Lisabon: University, 2014, s. 105. URL: http://www.smtda.net/images/Final_BOOK_OF_ABSTRACTS_SMTDA2014.pdf

10. Petrová Kafková, M. (2013). Šedivějící hodnoty? Aktivita jako dominantní způsob stárnutí. Brno: Masarykova univerzita. ISBN: 9788021063105.
11. Prevalence demence v Evropě. Česká Alzheimerovská společnost. 2015. http://www.alzheimer.cz/alzheimerova-choroba/vyskyt-demence/
12. Úmrtnost. Český statistický úřad: Ediční plán 2014. 2014, p. 39-47.http://www.czso.cz/csu/2014edicniplan.nsf/t/420031BBE5/$File/130069-14_6.pdf
13. WHO. World Health Organization. 2015. Dementia. http://www.who.int/mediacentre/factsheets/fs362/en/

Fair value of a Longevity Swap with Debt Value Adjustment

Valeria D'Amato[1], Luisa De Martino[2]

[1]Department of Economics and Statistics
University of Salerno, Italy
Email: vdamato@unisa.it

[2] University of Salerno, Italy
Email: l.demartino5@studenti.unisa.it

Abstract.
The longevity risk consists in the systematic deviations of the deaths' number from its expected value caused by improvements of mortality trend. This risk impacts ruinously on balance sheets of many holders, such as governments, pension providers, individuals and insurers. Transferring of the longevity risk to the capital market opens new opportunities, since it can become tradable.
The most popular instrument for managing longevity risk consists in longevity swap (bespoke o indexed). In a longevity swap, the pension fund agrees to make to the counterparty (insurer or bank) periodic payments based on agreed mortality assumptions. In return the counterparty pays to the pension fund regular floating flows based on either the pension plan realized survival rates either an agreed mortality index. These instruments involve counterparty risk, i.e. the risk that the other part of a contract will not live up to its obligations.
Nevertheless it is necessary to determine the price of longevity swaps in the most appropriate way, by considering an evaluation based on integrated risk analysis. In particular the counterparty risk should be taken into account as relevant risk, in order to guarantee the market transparency and standardization of longevity swap.
In recent years innovative instruments of collateralization have been developed such as Credit Value Adjustment (CVA), Bilateral CVA (BCVA) and Debt Value Adjustment (DVA). While CVA represents the price of the counterparty risk in the derivative contract computed assuming the counterparty may default prior the contract maturity and the investor is default risk-free and the BCVA is the cost of the counterparty risk considering its bilateral nature under the assumption that both party of the derivative contract may default prior the contract maturity, the DVA consists in the cost of the counterparty risk under the assumption that only the investor may default and his counterparty is default risk-free. In particular it is typically defined as the difference between the value of the derivative assuming the bank is default-risk free and the value reflecting default risk of the bank. In this contest we present a pricing formula for a fair valued longevity swap, since to do this it is necessary to consider market variables and the creditworthiness of both parties entering into the contract, as well as the valuation methodologies used by parties.
Typically a bank or insurance, which grants a longevity swaps to the pension funds, can be interested in including the cost of counterparty risk in a such instrument, by forcing the pension fund to pay a DVA.
In this context we propose a pricing model for a *completed* longevity swap, being a DVA included.
An empirical analysis is provided.

Keywords: Longevity Risk, Counterparty risk, Longevity swap, Debt Value Adjustment, Credit Value Adjustment, bilateral Credit Value Adjustment.

1. Introduction

With few exceptions, longevity has been increasing throughout the world during the last century. The emergence of increasing longevity is the result of substantial demographic changes over the last century. Particularly several studies show that the aging of the population is due to two firmly established and parallel trends: life expectancy is increasing and birth rates remain at historically low levels. Besides, the first half of the 20th century saw substantial reductions in early life mortality, while the second half century has shown significant improvements in mortality rates at older ages.
These demographic changes have caused an underestimation over the year of the like expectation and the mortality probabilities. This fact represents a financial risk so-called demographic risk.
The demographic risk is divided in two components: the insurance risk and the longevity risk. The insurance risk arises from accidental deviations of the number of deaths from its expected values, and it is a pooling risk, i.e. it can be mitigated by increasing the number of policies. The longevity risk derives from improvements in mortality trend, which determine systematic deviations of the number of deaths from its expected values (Di Lorenzo and Sibillo 2002).
The risk can be transferred from those who hold it, including individuals, governments, and private providers of retirement income, to insurers and banks through de-risking strategies as for instance buy-outs and buy-ins.
Nevertheless banks and insurers have a finite capacity to take longevity risk. They may pass longevity risk at the reinsurer offering a transaction similar to the longevity risk mentioned above. Currently there are many proponents of a capital market solutions where longevity risk can become tradable. So that this to occur, it is need transformed a non-financial risk in a financial asset. The market for the longevity-linked securities has developed substantially in recent years. The longevity-linked securities allow to the holders of the longevity risk (particularly the pension fund) of transferring longevity risk to the capital market. The longevity-linked securities are derivatives or bonds where the underlying is the trend of the mortality of the reference population.

The most popular longevity-linked instruments the longevity swap. In a longevity swap, the pension fund agrees to make to the counterparty (insurer or bank) periodic payments based on agreed upon mortality assumptions. In return the counterparty provides the pension fund with regular floating payment based on either the pension plan realized survival rates or an agreed upon mortality index.
Certainly, we must consider that with few exceptions, the longevity swap pricing formulas views so far pose an issue of opacity. Indeed we believe that for allowing an improvement of market liquidity and for achieving a standardized market where exchanging the longevity swap, we cannot exempt from considering the counterparty risk that this instrument may involve.
Counterparty risk consists in the risk that the entity with whom one has entered into a financial contract (the counterparty to the contract) will fail to fulfill their side of the contractual agreement, i.e. will default prior to expiration of a trade and will not therefore make the current and future payments required by the contract.
Then this risk is similar to other forms of credit risk in that the cause of economic loss is obligor’s default. There are, however, two features that set counterparty risk apart from more traditional forms of credit risk: the uncertainty of exposure and bilateral nature of credit risk (Pykhtin and Rosen 2009).
In addition in longevity swap counterparty risk exposure is large, since generally the longevity swaps have very long maturities.
We believe that if the counterparty risk is not considered in the pricing formula of the longevity swap, it could lead at an issue of mispricing.
In light of these considerations, we propose pricing formulas of the longevity swap that include the counterparty risk through innovative collateralization tools, i.e. the Debt Value Adjustment (from herein DVA).
The pricing formula of the longevity swap that involves the DVA allows to obtain a "completed" pricing formula; indeed, allows to the hedger of the longevity risk of reaching to a value of the swap that considers an important risk as counterparty risk, obtaining the riskiness value of the longevity swap.
The pricing formula of the longevity swap that involves the DVA is important for the development of a standardized and liquid market, where it is possible to exchange the longevity swaps. Indeed this pricing formula could be used by an authorities to guarantee the efficiency and transparency of the market.
The paper is organized as follows. The paper is organized as follows. In section 2 we develop formulas for pricing the fair value of longevity swap with Debt Value Adjustment. Section 3 illustrates the main outcomes of the empirical application.

2. Pricing a completed longevity swap: the case of debt value adjustment

According to Biffis et al (2014) the market value of the longevity swap’s payment that occurs at the maturity t=T at the inception is given by:

$$S_0 = nE^P\left[e^{-\int_0^T rtdt}\left(\frac{n-NT}{n} - P^N\right)\right] = nE^P[\,e^{-\int_0^T (rt+\mu t)dt} - B(0,T)P^N)] \qquad (2.1)$$

where P^N denotes the survival fixed rate referred to the maturity T; while $e^{-\int_0^T (rt+\mu t)dt}$ the risk neutral survival probability at the same maturity T referred to the homogeneous group of individuals, i.e. the estimation of the floating rate at the maturity T, where rt is the risk-free interest rate and μt is the mortality intensity; B(0,T) corresponds to the discounting factor; n denotes both the notional that the policyholders’ number that join to the pension fund. As regard the intensity of mortality and then the survival probability may be modeled using a stochastic mortality model.
By this equation it is obtained the survival fixed rate referred at the maturity T relying on the assumption that the expected value of each payment at each maturity of the longevity swap evaluated at time t=0 is equal to 0. Then the survival fixed rates are determined letting the above equation equal to 0 and solving for p^N:

$$p^N = \tilde{P}T + B(0,T)^{-1}Cov^{\tilde{\mathbb{P}}}\left(e^{-\int_0^T rtdt}, e^{-\int_0^T \mu tdt}\right) \qquad (2.2)$$

The survival fixed rate is given by the risk-neutral survival probability, $\tilde{P}T$, more the covariance between intensity of mortality and bond market returns.
This equation shows that if the intensity of mortality is uncorrelated with bond market returns, the latter term of the equation (2.2) is equal to 0 and the survival fixed rates associated to different maturities involve only the survival probabilities $\{\tilde{P}T_i\}$ referred at the same maturities. Then the survival fixed rates referred each maturity are determined basing only on the mortality forecasts made at the inception of the longevity swap.
The pension fund cannot reach to a complete assessment of the longevity swap namely based on integrated risk analysis, unless it is not taken into account the counterparty risk on the pricing formula. In this regard we can use innovative collateralization tools such as Debt Value Adjustment (DVA), indeed only the consideration of the counterparty risk allows to represent an internal management model which take into account an integrated analysis of risks.

In D'Amato and De Martino (2014a, 2014b) formulas of evaluation of the longevity swap by including the unilateral and bilateral counterparty risks from the point of view of the pension fund through the credit value adjustment (CVA) and bilateral credit value adjustment (BCVA) have been proposed, being the CVA the market price of counterparty risk on a contract obtained by the risk neutral expectation of the loss that could occur for the counterparty default over the term of the contract and the BCVA the cost of the counterparty risk under the assumption that both counterparties could default.
However it could result necessary to assess the present value of the longevity swap considering the impact of the risk that the pension fund could default and his counterparty is default free, then to consider the counterparty risk from the point of view of the pension fund's counterparty.
In this way a such evaluation of the longevity swap could allow to the pension fund of analyzing the impact of the own default on the longevity swap value and identifying the entity of the collateral that could be required from the counterparty institution.
For this kind of evaluation we can use the DVA as a spread, namely the market price of counterparty risk on a contract obtained by the risk neutral expectation of the gain that could occur for the pension fund default over the term of the contract weighted with the risk-neutral probability of the own default.
The DVA is analogous to the CVA and is the price of counterparty credit risk from the perspective of the counterparty, i.e., the price of the risk that the investor defaults before maturity of a derivative contract and fails to full his obligations to the counterparty. Then the DVA for the pension fund is the CVA for his counterparty.
The unilateral DVA as a stand-alone value is given by:

$$UDVA = (1 - \bar{\bar{\delta}}_I) \sum_{i=1}^{T} B(t_i) NEE(t_i) q_F(t_i, t_{i-1}) \qquad (2.3)$$

$q_I(t_i, t_{i-1})$ denotes the default probability of the institution; $\bar{\delta}_I$ denotes the recovery of the institution; while $NEE(t_i)$ denotes the negative expected exposure, i.e. the EE from the point of view of the counterparty with the difference that the $NEE(t_i)$ is a negative value.

The unilateral DVA as a credit spread is given by the following equation:

$$UDVA_{as\ a\ spread} = X_F^{CDS} \times ENE \qquad (2.4)$$

where X_F^{CDS} is the periodic premium paid by the investor that enter into the credit default swap to cover his counterparty risk exposure and ENE is the expected negative exposure.
The DVA as a stand-alone value and as a credit spread unlike the CVA are negative values.
The pension fund can evaluate approximately the P&L (profit and loss) impact of the longevity swap considering the cost of counterparty risk under the assumption that the pension fund could default and his counterparty is default free, simply subtracting the DVA as a running spread at the expectation under the risk neutral measure of net rate that he will receive, i.e. the different among the floating rate and the fixed rate that occurs at each maturity, as reported by the following equation:

$$S_0 = n\left\{E^{\widetilde{\mathbb{P}}}\left[(e^{-\int_0^T (rt+\mu t)dt} - B(0,T)P^D\right] - DVA_{Spread}\right\} \qquad (2.5)$$

However as it is possible to note the consideration of the cost of the own default through of the DVA as a spread increases the risky market value of the longevity swap, since a negative value, the DVA, is subtracted to the expectation under $\widetilde{\mathbb{P}}$ of the net rate that could receive the pension fund. Indeed if the pension fund defaults and the present value of the longevity swap form his point of view is negative, he will make an gain. This because the pension fund should pay an amount, that will not be paid in case of own default. The formula (2.5) might be used by the pension fund for self-rating the impact of the own default on the present value of the longevity swap in such a way that he can check the consistency of the collateral potentially required by institutions.
In addition the DVA could be understood as a charged to apply to the pension fund for covering the cost of the counterparty risk from the point of view of the institution that grants the longevity swap. Indeed an institution that grants the longevity swap to the pension fund, having major bargaining power and higher credit quality than the pension fund, could require that the longevity swap is completed of collateral and the collateral could be an DVA as a spread.
Basing on this consideration, you can determine the survival fixed rate of a longevity swap that includes as collateral a DVA simply putting equal to zero the equation (1.4) and solving for P^D. In this way you obtain:

$$P^D = \tilde{P}T + B(0,T)^{-1} Cov^{\widetilde{\mathbb{P}}}\left(e^{-\int_0^T rtdt}, e^{-\int_0^T \mu tdt}\right) - DVA_{spread}\, B(0,T)^{-1} \qquad (2.6)$$

In this case the fixed rates of the longevity swap depend not only on the survival probabilities and the covariance among the intensity of the mortality and the bond market returns, but also on the market value of the cost of counterparty risk under the assumption that the pension fund could default and his counterparty is default free.
To determine the survival fixed rates considering the charge to cover the counterparty exposure of the institution trough the DVA as a spread involves that the pension fund pay at each maturity a net rates higher than those that he should pay if it is not considered the cost of the own default.
Then this higher amount that the pension fund overpays could be accounted by his counterparty to cover the potential losses that could arise in the event of pension fund default.

3. Numerical Applications

We consider a pension fund with 1500 policyholders from a group composed by 65-67 years old Italian individuals. The pension consists of a payment of €25000 each year to each policyholder alive. For covering the longevity risk exposure the pension fund could enter in the fixed side of a longevity swap.
Then let us suppose that the pension fund enter in a 20-year indemnity longevity swap where the counterparty is an insurer.
Basing on this kind of contract the pension fund agrees to pay a series of fixed payment obtained multiplying the survival fixed rates (or pre-determined) to the notional at each maturity, that occurs every July 10th from 2015 to 2035. Whilst in return the pension fund receives by the insurer a series of floating payments obtain multiplying effective survival rate of the policyholder to the notional at the same maturity date. You let us assume that the notional of the longevity swap is €37,500,000.
The first issue to face up, which involves an high cost and time intensity analysis on the pension plan's book, is of estimating the floating rate and defining the fixed rates also called forward survival rate.
For estimating the floating rates of the longevity swap it is needed simply to calculate the contingent probabilities that an Italian individual group aged between 65-67 today survives every year up to the end of the next 20 years. These probabilities are obtained with the Lee and Carter model (1992) and are reported in the table 1.

Table 1 – Best estimation of floating rates, Lee and Carter model

	Floating rate
P65-67(2014)	0,990877167
P65-67(2014:2015)	0,981010215
P65-67(2014:2016)	0,970360249
P65-67(2014:2017)	0,958991816
P65-67(2014:2018)	0,94664446
p65-67 (2014:2019)	0,933217517
p65-67 (2014:2020)	0,918457062
p65-67 (2014:2021)	0,902319878
p65-67 (2014:2022)	0,884907835
p65-67 (2014:2023)	0,866054588
p65-67 (2014:2024)	0,845696765
p65-67 (2014:2025)	0,823516197
p65-67 (2014:2026)	0,799495083
p65-67 (2014:2027)	0,772747317
p65-67 (2014:2028)	0,742871152
p65-67 (2014:2029)	0,709872659
p65-67 (2014:2030)	0,674377512
p65-67 (2014:2031)	0,636713078
p65-67 (2014:2032)	0,596593789
p65-67 (2014:2033)	0,553913895
p65-67 (2014:2034)	0,508583842

For setting the survival fixed rates referred at each maturity we use the same forecast of the mortality trend, but we include into the survival rates also the risk premium that the insurer requires for granting the longevity swap to the pension fund. Indeed, an insurer o bank enter into the floating side of longevity swap only if the longevity risk assumed is remunerated with adequate risk premium.
Then supposing that the contract provides a risk premium given by an additional mortality improvement of 0.5% and by the 95% of the base mortality table, i.e. the mortality table of the year 2010.
The table 2 reports the contingent survival probabilities of the Italian population aged between 65-67 years today that include also the risk premium for the longevity risk.

Table 2 – Fixed rates, Lee and Cartel model

	Fixed rate
p̌65-67(2014)	0,991507
p̌65-67(2014:2015)	0,982363
p̌65-67(2014:2016)	0,972535
p̌65-67 (2014:2017)	0,962089
p̌65-67(2014:2018)	0,950789
p̌65-67 (2014:2019)	0,93855
p̌65-67(2014:2020)	0,925145
p̌65-67(2014:2021)	0,910541
p̌65-67(2014:2022)	0,894834
p̌65-67(2014:2023)	0,877878
p̌65-67(2014:2024)	0,859618
p̌65-67(2014:2025)	0,839768
p̌65-67(2014:2026)	0,81831
p̌65-67(2014:2027)	0,794448
p̌65-67(2014:2028)	0,76781
p̌65-67(2014:2029)	0,738378
p̌65-67(2014:2030)	0,706678
p̌65-67(2014:2031)	0,67296
p̌65-67(2014:2032)	0,636916
p̌65-67(2014:2033)	0,598378
p̌65-67(2014:2034)	0,557166

Once estimated the floating rates and set the survival fixed rates before of calculating the expected present value of the longevity swap it is necessary to choice the discounting rate. In this regard considering that the reference population is Italian we can use as discounting rate the yield of the Italian bearing coupon bonds i.e. BTP issued at 14/07/2014 with maturity at 15 years, i.e. 3.5%.
Table 1.3 shows the discounting factors for each maturity of the longevity swap.

Table 3 – Present value of €1, discounted to the n^{th} year

DATE	Discount Factor
10/07/2014	1
10/07/2015	0,966183575
10/07/2016	0,9335107
10/07/2017	0,901942706
10/07/2018	0,871442228
10/07/2019	0,841973167
10/07/2020	0,813500644
10/07/2021	0,785990961
10/07/2022	0,759411556
10/07/2023	0,733730972
10/07/2024	0,708918814
10/07/2025	0,684945714
10/07/2026	0,661783298
10/07/2027	0,639404153
10/07/2028	0,61778179
10/07/2029	0,596890619
10/07/2030	0,576705912
10/07/2031	0,557203779
10/07/2032	0,53836114
10/07/2033	0,52015569
10/07/2034	0,502565884
10/07/2035	0,485570903

Now it is possible to build up the cash flows of the longevity swap and to evaluate the expected present value of the longevity swap as shown by the table 4.

Table 4 – Evaluation of the indemnity longevity swap

Year	Payment date	Fixed rate	Floating rate	Difference between floating rate and fixed rate	Net payment	Present value Net Payment
2014	10/07/2015	99,15%	99,09%	-0,06%	-23636,27	-22836,98
2015	10/07/2016	98,24%	98,10%	-0,14%	-50725,37	-47352,68
2016	10/07/2017	97,25%	97,04%	-0,22%	-81564,10	-73566,14
2017	10/07/2018	96,21%	95,90%	-0,31%	-116139,79	-101209,12
2018	10/07/2019	95,08%	94,66%	-0,41%	-155427,64	-130865,90
2019	10/07/2020	93,85%	93,32%	-0,53%	-199956,19	-162664,49
2020	10/07/2021	92,51%	91,85%	-0,67%	-250786,33	-197115,78
2021	10/07/2022	91,05%	90,23%	-0,82%	-308280,17	-234111,52
2022	10/07/2023	89,48%	88,49%	-0,99%	-372236,79	-273121,66
2023	10/07/2024	87,79%	86,61%	-1,18%	-443385,22	-314324,12
2024	10/07/2025	85,96%	84,57%	-1,39%	-522035,33	-357565,86
2025	10/07/2026	83,98%	82,35%	-1,63%	-609431,55	-403311,62
2026	10/07/2027	81,83%	79,95%	-1,88%	-705562,11	-451139,34
2027	10/07/2028	79,44%	77,27%	-2,17%	-813766,90	-502730,37
2028	10/07/2029	76,78%	74,29%	-2,49%	-935192,10	-558207,39
2029	10/07/2030	73,84%	70,99%	-2,85%	-1068954,66	-616472,47
2030	10/07/2031	70,67%	67,44%	-3,23%	-1211262,22	-674919,88
2031	10/07/2032	67,30%	63,67%	-3,62%	-1359259,61	-731772,55
2032	10/07/2033	63,69%	59,66%	-4,03%	-1512086,18	-786520,23
2033	10/07/2034	59,84%	55,39%	-4,45%	-1667404,24	-837980,49
2034	10/07/2035	55,72%	50,86%	-4,86%	-1821840,08	-884632,53
Present value of the swap					-14228932,84	**-8362421,2**

The first column reports the years of reference of the survival rates; the second column the maturity date. In particular the payment is achieved the next year to the reference year of the survival rates. This is due by the fact that since the floating rates paid by the insurer is the effective survival rates of the reference population we could know these rates only if the reference year is concluded. The third and the fourth column report respectively the survival fixed rates and the best estimation of the floating rates. While the fifth column shows the net rates received by the pension fund at each maturity, namely the difference between the best estimation of the floating rates and the survival fixed rates. How it is possible to note the net rates are negatives, that’s why the best estimation of the floating rates and the survival fixed rates are determined basing on the same mortality forecast, however the survival fixed rates include also the risk premium for the longevity risk required by the insurer.
The net rates multiplied for the notional allow to obtain the net payments for each maturity as shown in the sixth column. The last column shows the present value of each payment which sum gives the expected present value of the longevity swap from the point of view of the pension fund. This value is about €-8,362,000 for the presence of the risk premium. The following figure 1 shows the difference between the floating and the fixed rates underlining the entity of the risk premium

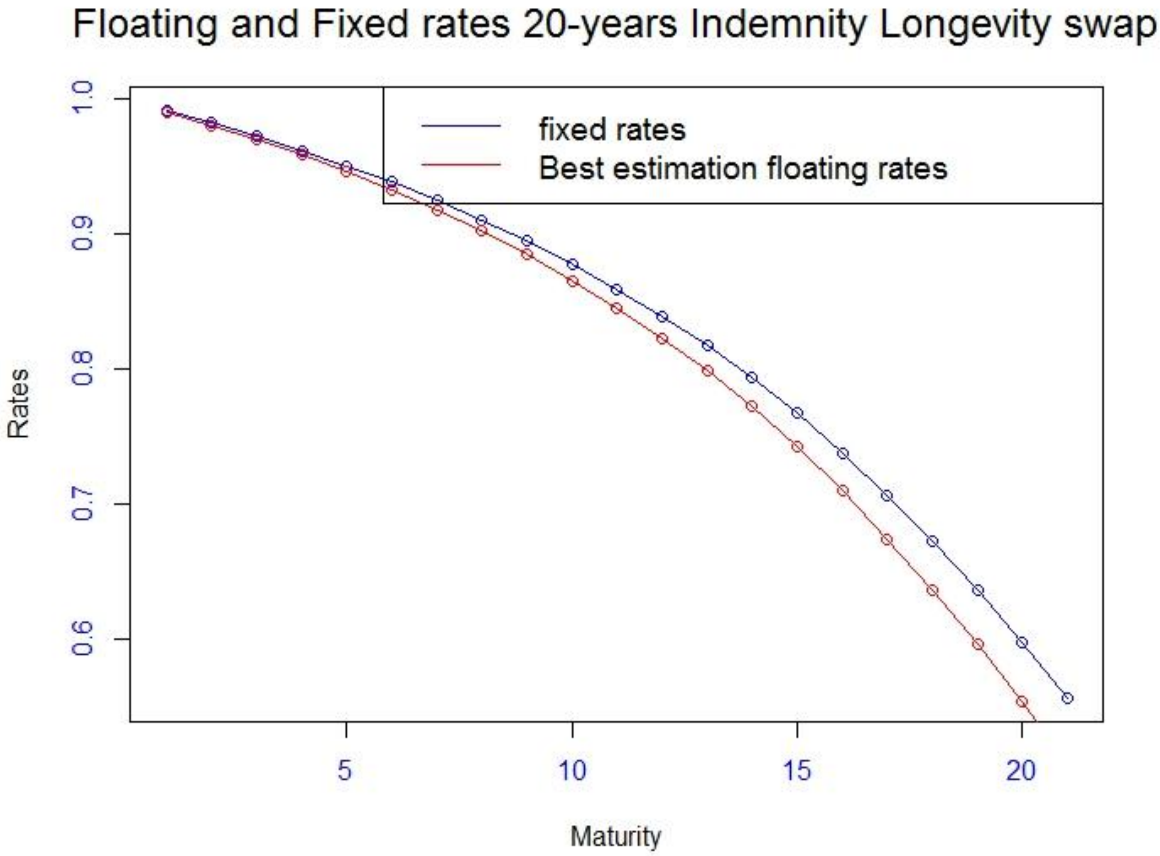

Figure 1 – Trend of the forward survival rate and best estimation of the survival rate

As it is possible to note, this difference increases over the time. This because over the time increases the uncertainty of the prediction, therefore also the entity of the risk premium increases over the time.

Certainly it must be considered that the evaluation of the indemnity longevity swap, as well as made, is incomplete. Indeed this evaluation do not consider the counterparty risk. In D'Amato and De Martino (2014a, 2014b) we have assessed the longevity swap considering the impact of the unilateral counterparty risk of the pension fund and the bilateral nature of the counterparty risk through innovative collateralization tools such as CVA and BCVA.

However we have not considered the impact of the counterparty risk under the assumption that the pension fund could default and the insurer is default-free. For measuring the influence on the expected present value of the longevity swap of the own default we iclude another innovative collateralization tool such as DVA as a spread. As aforementioned the DVA is the market price of the counterparty risk on the contract obtained considering the risk neutral expectation of the gains that could derive from the own default over the term of the contract.

It is intuitive that the consideration of the impact of the own default on longevity swap from the point of view of the pension fund increases the expected present value of this contract. Indeed if we consider that the pension fund fails and the value of the contract is negative, it will get a gain, since it will not perform the future payment.

As pointed out in section 2 the DVA as a spread is given by the following equation:

$$UDVA_{as\ a\ spread} = X_F^{CDS} \times ENE$$

where X_F^{CDS} is the periodic premium paid by the insurer that enter into the credit default swap to cover his counterparty risk exposure and ENE is the expected negative exposure. The $UDVA_{as\ a\ spread}$is always negative.

We suppose that the X_F^{CDS} is equal to 3% and the ENE is equal to -4% we have a DVA as a spread equal to:

$$UDVA_{spread} = -4\% \times 3\% = -0.12\%$$

Once calculated the $UDVA_{as\ a\ spread}$ we can assess the expected present value of the longevity swap considering the cost of the default of the pension fund simply subtracting this spread to the floating rates received by the pension fund (or adding to fixed rates paid by the pension fund) for each maturity.

The table 5 shows the cash flows of the longevity swap for the pension fund considering the impact of the own default.

Table 5 - Valuing the indemnity longevity swap with DVA as a spread

Year	Payment date	Fixed rate	Floating rate	Floating rate - dva spread	Difference between floating rate(including dva spread) and fixed rate	Net payment	Present value Net Payment, considering counterparty risk
2014	10/07/2015	0,991507	0,990877167	99,21%	0,057%	21363,7257	20641,281
2015	10/07/2016	0,982363	0,981010215	98,22%	-0,015%	-5725,3742	-5344,6981
2016	10/07/2017	0,972535	0,970360249	97,16%	-0,098%	-36564,097	-32978,721
2017	10/07/2018	0,962089	0,958991816	96,02%	-0,190%	-71139,79	-61994,217
2018	10/07/2019	0,950789	0,94664446	94,78%	-0,294%	-110427,64	-92977,106
2019	10/07/2020	0,93855	0,933217517	93,44%	-0,413%	-154956,19	-126056,96
2020	10/07/2021	0,925145	0,918457062	91,97%	-0,549%	-205786,33	-161746,19
2021	10/07/2022	0,910541	0,902319878	90,35%	-0,702%	-263280,17	-199938
2022	10/07/2023	0,894834	0,884907835	88,61%	-0,873%	-327236,79	-240103,77
2023	10/07/2024	0,877878	0,866054588	86,73%	-1,062%	-398385,22	-282422,77
2024	10/07/2025	0,859618	0,845696765	84,69%	-1,272%	-477035,33	-326743,3
2025	10/07/2026	0,839768	0,823516197	82,47%	-1,505%	-564431,55	-373531,37
2026	10/07/2027	0,81831	0,799495083	80,07%	-1,761%	-660562,11	-422366,16
2027	10/07/2028	0,794448	0,772747317	77,39%	-2,050%	-768766,9	-474930,19
2028	10/07/2029	0,76781	0,742871152	74,41%	-2,374%	-890192,1	-531347,31
2029	10/07/2030	0,738378	0,709872659	71,11%	-2,731%	-1023954,7	-590520,7
2030	10/07/2031	0,706678	0,674377512	67,56%	-3,110%	-1166262,2	-649845,71
2031	10/07/2032	0,67296	0,636713078	63,79%	-3,505%	-1314259,6	-707546,3
2032	10/07/2033	0,636916	0,596593789	59,78%	-3,912%	-1467086,2	-763113,22
2033	10/07/2034	0,598378	0,553913895	55,51%	-4,326%	-1622404,2	-815365,02
2034	10/07/2035	0,557166	0,508583842	50,98%	-4,738%	-1776840,1	-862781,84
Present value of the swap						-13283933	**-7701012,3**

As it is possible to note from this cash flows, we have added a column where we have reported the floating rates minus the $UDVA_{as\ a\ spread}$.

Then the $UDVA_{as\ a\ spread}$, that is always negative, is subtracted to the rates received by the pension fund, increasing the expected present value of the longevity swap for the pension fund.

Indeed, the present value of the swap for the pension fund obtained considering the impact of the own default is again a negative number but is higher than the expected present value of the free risk longevity swap.

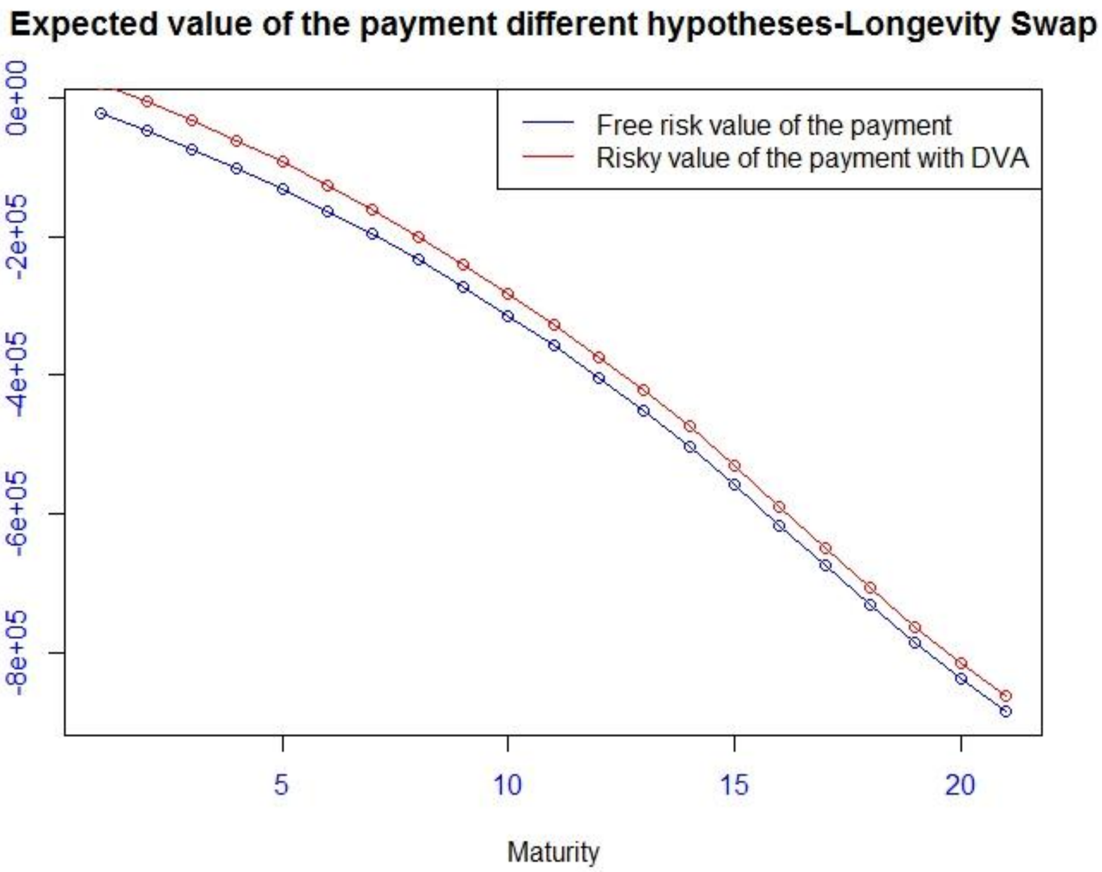

Figure 2 - Expected value of the payment in the different hypotheses

This graph underlines the difference between the free risk value of the longevity swap's payment and its risky market value obtained under the assumption that the pension fund could default and his counterparty is default free.
Finally we can affirm that the examination of the own default on the longevity swap through the DVA as a spread from the point of view of the pension fund involves that the risky present value of the contract is higher than its free risk value.
However the counterparty institution which grants the longevity swap to the pension fund may require as collateral the DVA. In this latter case, also if in the pricing formulas will appear the DVA, the present value of the swap results decreased by the present of this kind of collateral.

Conclusion

In the previous sections, we have considered the longevity risk and the most popular instrument for transferring this risk to the capital market, i.e. the longevity swap. Although this instrument allows to transfer this important risk which impact the balance sheet of many stakeholders, it involves another risk such as counterparty risk.
Without an integrated analysis risk it may occur problems of mispricing. In particular it could be important an evaluation of the longevity swap that consider the impact of the own default from the point of view of the pension found, allowing at this latter of identifying the entity of the collateral that could be required from the counterparty institution and, exceptionally, of reaching an fair valued longevity swap formula.
We have proposed an pricing formula that includes an innovative collateralization tools, i.e. the DVA, for obtaining the fair value of the longevity swap.
The longevity swap valued with the formula of pricing proposed, namely that include the impact of the own default for the pension fund, results greater than an longevity swap valued with an classic formula, that then does not include the cost of the own default.
In addition we have proposed a formula to determine the fixed rates of the longevity swap including again DVA as a spread, allowing to the pension fund to analyze the impact of the collateral that may be required from the counterparty of the pension fund.
Further researches will be on the comparison between different pricing formulas that include different collateralization tools such as CVA, BCVA and DVA.

References

Biffis, E. Blake, D. Pitotti, L. Sun, A. (2014) The cost of the counterparty risk and collateralization in longevity swaps. The journal of risk and insurance, vol. 9999, 1-33.

D'Amato, V. De Martino, L. (2014a) Pricing collateralized longevity swap in Credit Value Adjustment setting, presented to Perspectives on Actuarial Risks in Talks of Young-Researchers- PARTY 2015.

D'Amato, V. De Martino, L. (2014b) Marking to Market Pricing for a Bespoke collateralized Longevity Swap, under review.

Di Lorenzo, E. Sibillo, M., (2002) Longevity Risk: Measurement and Application Perspectives.

Gregory, J. (2012) Counterparty credit risk and credit value adjustment.

Lee, R.D., Carter, L.R., (1992), Modeling and Forecasting U.S. Mortality, Journal of the American Statistical Association 87, 659-671.

Pykhtin ì, M. Rosen, D (2009) Pricing counterparty risk at trade level and CVA allocations

Detection of similar behaviors and abnormal segments in time series

Christian Derquenne

Electricité de France, Research and Development, 1 avenue du Général de Gaulle, 92141 Clamart Cedex, France
(E-mail: christian.derquenne@edf.fr)

Abstract. The objectives of time series analysis are many: forecast their future behavior, understand how the response is built using predictors, synthesize information from several time series, detect breaks behavior or search similar or abnormal time intervals. This paper is placed in the context of the latter objective. For this, we use a method of segmentation of time series that we have developed. This method provides constant, increasing or decreasing linear segments. Then we introduce a method of ascending hierarchical clustering of segments using a stopping rule to determine the number of clusters. This criterion is based on statistical tests comparing the coefficients associated with the segments. Therefore, each cluster will contain similar time intervals while those which include a single element may correspond to abnormal periods relative to the general behavior of the time series. Tests of multiple comparisons are used to decide if these clusters are atypical compared to others. This approach is applied to data coming from the energy management. Finally, we propose future directions based on the approach introduced previously.
Keywords: Time series, segmentation, clustering.

1 Context and issues

The time series are decomposed into several types of changes: trend, seasonality, volatility and noise. They may be more or less regular according to the application domain. Behavioral changes that characterize these series are mainly of several types: peak (price of energy in tense situation, but on a very short period), jumps in level or trend (data stream), jumps variability (yield of the FTSE 100). Modeling of these series is very delicate and requires a lot of experience in the application domain. It may be interesting to detect changes in behavior for many applications in the pre-treatment: construction of sub-models in each segment, stationnarized series using segmentation, building of symbolic curves to achieve a clustering of curves, modeling of multivariate time series, etc. Many segmentation methods in Arlot [1] , Guédon [9], Lavielle and Teyssière [10] have been and are developed to address various problems in economics, finance, human sequencing, meteorology, energy management, etc. Most of these methods rely on the use of dynamic programming to reduce drastically the number of possible segmentations because it would obviously be

16^{th} *ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece*

totally illusory to calculate them all. Indeed, the number of segmentations for a serie of length T and M fixed number of segments is $\binom{T-1}{M-1}$ whereas the set of all segments $M = 1, T$, the total number of segmentations increases to 2^{T-1}. The complexity of these algorithms is in general $O(T^2)$. These methods to detect break points are designed to solve three problems: (i) detecting a change in the mean, with a constant variance, (ii) detecting the change in variance with a constant mean (iii) detecting changes in the overall distribution of the phenomenon, without distinguishing changes in level, variability and distribution errors.

We introduced a method in Derquenne [3] which not only reduces the complexity compared to other methods, but above all proposes solutions segmentation of the series containing segments which are constant , which increase or decrease with different dispersions. Our method is original in its approach because it offers a decision support for time series, step by step. It contains two main phases: data preparation to obtain a first segmentation of data and modeling of segments based on a Gaussian heteroskedastic linear model by successive adaptations. Each of the two phases is repeated a few times depending on the degree of smoothing applied to the data. The degree of smoothing can vary from 1 to T theory. It corresponds to the number of observations included in moving median used in the phase of data preparation. The empirical complexity is $O(T\sqrt{T})$ and the theoretical complexity is $O(T^2)$. This method has been improved by a better consideration of the variability of the data in Derquenne [4] and through a meta-segmentation approach in Derquenne [5], selecting the best segments from different degrees of smoothing j available. This method has been tested on many series and has provided encouraging results on both simulated data to assess the quality of reconstruction of the series: detection of breakpoints and modeling segments, but mainly on real data, especially in the field of training in energy market prices.

If this method of segmentation and many others used to meet most of the following questions: "Are there breaks behavior and how to detect them ?" it is also important to address other issues associated with the problems encountered in different fields of applications such as: (i) Are there similar behaviors between these series? in particular, are there common failures? (ii) Are there common features groups and/or different lines? (iii) Are there any unobservable leverage of one or more temporal phenomena to explain? (iv) Is a response series systematically explained by the same inputs along the entire length of the series? (v) Which are weights of these inputs? (vi) Are there similar behaviors in the same time series? (vii) Are there atypical behaviors in the same time series?

We have proposed approaches to address (i) and (ii) in Derquenne [6], (iii) in Derquenne [7] and, (iv) and (v) in Derquenne [8]. The objective of this paper is to introduce a method to answer to (vi) and (vii). Section 2 below, formalizes this approach, and Section 3 is devoted to an application on the prices of the energy market, and finally the last section concludes on inputs, potential improvements of the proposed approach, as well as future researches and applications based on the segmentation of time series.

2 Detection of similar and atypical behaviors

In case of irregular time series, stock prices (energy market, CAC40, FTSE100, GDP, etc.), it can be interesting to detect similar intervals time (segments) where the response follows same behavior (fig. 1.a). The proposed solution is based on clustering of segments $\tau_1, ..., \tau_m, ..., \tau_M$. Indeed, a cluster containing several segments is related to a similar behavior of parts of time series. On the other hand, if we analyse regular times series (fig. 1.b), such as temperature curve, detection of atypical segments can be fruitful to identify a problem on the data. The detection process is based on statistical tests of comparison of means.

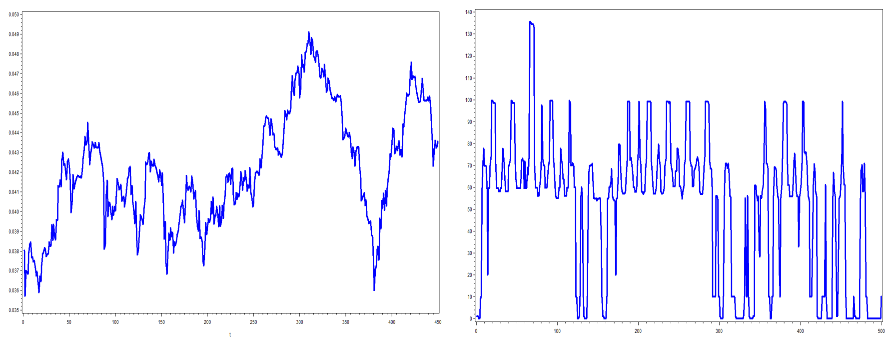

Fig. 1. (a) Irregular time series (b) Regular time series

2.1 The model and its inference

Let's be a time series $(Y_t)_{t=1,T}$, we assume that it decomposes according to the Gaussian heteroskedastic linear model (or variance components) [11] as follows:

$$Y_t = \sum_{m=1}^{M} (\beta_0^{(m)} + \beta_1^{(m)} t + \sigma_m \epsilon_t) 1_{[t \in \tau_m]} \tag{1}$$

where $(\beta_0^{(m)}, \beta_1^{(m)}) \in \mathbb{R}^2$ and $\sigma_m > 0$ are respectively the parameters of level, slope and dispersion for the segment τ_m and ϵ_t follows a standard Normal. This model has a structure of piecewise regression.

The number of observations per segment τ_m is denoted n_m with $\sum_{m=1}^{M} n_m = T$. Each segment τ_m contains the set of values: Y_t for $t = U_{m-1} + 1$ to U_m, where $U_m = U_{m-1} + n_m$, finally $U_M = T$. There are so $3M$ parameters to be estimated, knowing the number of segments M is unknown. The segmentation approach proposed in Derquenne [3], [4] and [5] is entirely unsupervised. Finally, to estimate the Gaussian heteroskedastic linear model we have used the REstricted Maximum Likelihood estimator (REML).

2.2 Detection of similar segments by clustering method

The proposed method is based on a hierachical ascending clustering controlled by a statistical stopping rule. Indeed, a statistical test of similarity of clusters will allow to aggregate or not two groups of segments. The process is as follows.

First step: M segments in M clusters

The dissimilarity d_{kl} between two clusters $G_k = \{\tau_k\}$ and $G_l = \{\tau_l\}$ (only one segment by cluster for the first step) is calculated with the p-value of statistical test associated to the comparison of regression coefficients by couple of segments. A preliminary test is necessary to verify the homoskedasticity of residual variances associated to two segments τ_l and τ_k.

$H_0 : \sigma_k^2 = \sigma_l^2 \ vs \ H_1 : \sigma_k^2 \neq \sigma_l^2$.

The statistics used is:

$$F_{obs} = \hat{\sigma}_k^2 / \hat{\sigma}_l^2 \tag{2}$$

where $\hat{\sigma}_k^2$ and $\hat{\sigma}_l^2$ are respectively the estimated residual variances associated to each segments τ_k and τ_l. Under H_0, (2) is distributed as a Fisher distribution $F(n_k - 2, n_l - 2)$.

If the p-value obtained is less than α fixed, for instance: 0.05 or 0.10, the error variances of τ_k and τ_l are different.

There are two tests to aggregate two clusters. The first one compares two segments with slope and intercept coefficients, whereas the second one tests only the equality of intercept coefficients of two segments.

For the first test, the hypothesis of equality of regression coefficients are as follows:

$H_0 : \left(\beta_0^{(k)} = \beta_0^{(l)}\right) \ vs \ H_1 : \left(\beta_0^{(k)} \neq \beta_0^{(l)}\right)$

$H_0 : \left(\beta_1^{(k)} = \beta_1^{(l)}\right) \ vs \ H_1 : \left(\beta_1^{(k)} \neq \beta_1^{(l)}\right)$

In addition, this test depends on result of the previous homoskedasticity test (2). If the null hypothesis of this last one is rejected then the statistic of comparison of regression coefficients is:

$$t_1^{(k,l)} = \left[\beta_1^{(k)} - \beta_1^{(l)}\right] \left[\frac{\hat{\sigma}_k^2}{n_k S_k^2} + \frac{\hat{\sigma}_l^2}{n_l S_l^2}\right]^{-1/2} \tag{3}$$

and

$$t_0^{(k,l)} = \left[(\beta_0^{(k)} - \beta_0^{(l)}\right] \left[\frac{\hat{\sigma}_k^2}{n_k} + \frac{\hat{\sigma}_l^2}{n_l}\right]^{-1/2} \tag{4}$$

where S_k^2 and S_l^2 are respectively the variance of the centered times index of the segments τ_k and τ_l.

If the homoskedasticity test is kept then $\hat{\sigma}_k^2$ and $\hat{\sigma}_l^2$ are replaced by $\hat{\sigma_{kl}}^2$.

Under H_0, $t_1^{(k,l)}$ and $t_0^{(k,l)}$ follow a Student distribution with $n_k + n_l - 4$ degrees of freedom. If we suppose that there are M_1 segments which have slope and intercept coefficients, then this test is made for all couple of segments. We obtain $M_1(M_1-1)/2$ p-values: $\left(p_0^{(k,l)}, p_1^{(k,l)}\right)$.

For the second test, there are M_0 segments which are only an intercept coefficient, the hypothesis are the following:

$H_0 : \left(\beta_0^{(i)} = \beta_0^{(j)}\right)$ vs $H_1 : \left(\beta_0^{(i)} \neq \beta_0^{(j)}\right)$

and the statistic is:

$$t_{\bar{y}}^{(i,j)} = [\bar{y}_i - \bar{y}_j]\left[\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}\right]^{-1/2} \tag{5}$$

Under null hypothesis, this statistic follows a Student distribution with n_i+n_j-2 degrees of freedom. As for the previous test, $\hat{\sigma}_i^2$ and $\hat{\sigma}_j^2$ are replaced by $\hat{\sigma_{ij}}^2$, if the homoskedasticity is not rejected. In addition, we obtain $M_0(M_0-1)/2$ p-values.

Remarks: $M = M_0 + M_1$.

The stopping rule is the following, if

$$p_{\min}^{(1)} = \min_{(k,l),(i,j)}\left[\min\left(p_0^{(k,l)}, p_1^{(k,l)}\right), p_{\bar{y}}^{(i,j)}\right] > \alpha \tag{6}$$

then the two associated segments τ_k and τ_l (or τ_i and τ_j) are aggregated to constitute a cluster $G_{(1)}$. Finally there are $M-1$ clusters. However, if H_0 is rejected at this first step, the M clusters (M segments) are not separable.

Second step: $M-2$ clusters with one segment in each one and one cluster with 2 segments

The same process of statistical tests is applied on the $M-2$ clusters, but only $G_{(1)}$ is compared to M_1-2 segments if this segment has the slope and intercept regression coefficients or to M_0-2 segments, if $G_{(1)}$ has only an intercept parameter. It is an efficient way to reduce the time computing. The stopping rule based on the minimum of p-values (6) is used to decide if the process of aggregation of clusters continues or not. If the process continues then we have two cases. For the first one, we have two clusters containing two segments each ones $G_{(1)}$ (coming from first step) and $G_{(2)}$ (coming from second step) and $M-4$ clusters owning one segment each one. For the second case, we have one cluster which contains three segments denoted: $G_{(2)} = \left(G_{(1)}, \tau_k\right)$ and $M-3$ clusters having one segment each one.

The following steps of aggregation of clusters

This process continues while the minimum of p-values is greater than a fixed α. The more α is small, the more number of clusters is small. Finally, we obtain $\tilde{M}$ clusters of segments $\left(G_{(1)}, ..., G_{(m)}, ..., G_{(\tilde{M})}\right)$ which are similar in terms of behavior.

2.3 Detection of abnormal segments

We define an abnormal segment as a statistically far from the other segments.

The process of detection of abnormal segments is the following. We compare each segment τ_k to the $M_1 - 1$ or $M_0 - 1$ other segments depending on their structure (slope and intercept coefficients or only intercept coefficient). This comparison is made with the tests introduced in subsection 2.2. Then for each constant segment for example, we obtain $M_0 - 1$ p-values, then we apply a multiple test approach. For the segment τ_k, there are $M_0 - 1$ null hypothesis to test:

For $H_{(0,l)} : \left(\beta_0^{(k)} = \beta_0^{(l)}\right)$ vs $H_1 : \left(\beta_0^{(k)} \neq \beta_0^{(l)}\right)$; for $l = 1$ to $M_0 - 1$ and $l \neq k$.

For each test l, if p-value$\leq \alpha$, then $H_{0,l}$ is rejected.

When we use a multiple test approach, there are approximately $100\alpha\%$ of false positive results, then a simple rule to reject overall the null hypothesis is that the observed percentage of rejection is greater than $100\alpha\%$. The mean rate of false positive penalized introduced by Benjamini et al. [2] can be used. These authors show that the mean rate of false positive is approximatively equal to $\pi_0\alpha$, where π_0 represents the expected proportion of true null hypothesis. These authors give an estimation of the number of rejected tests, such as: $\hat{k} = \max\{k : p_{(k)} \leq \alpha k/m\}$, where $p_{(k)}$ is the k^{th} p-value in increasing order and m is the number of tests. Then the estimated proportion of rejected null hypothesis is equal to $\bar{\pi}(\hat{k}) = \hat{k}/m$. Then it will be interesting to compare the evolution of proportion of observed rejected tests to the evolution of proportion of expected rejected tests. Whenever the observed proportion is greater than the expected proportion, then intersection value corresponds to significant level of multiple test.

The more intersection value is high, the more proportion of observed rejected tests is high. In other words, the more intersection value is high, the more the segment is abnormal.

3 Application on energy management data

We apply our methodology on two databases of energy management: market prices and marginal costs. The name of these data is not given for confidential reasons.

3.1 Detection of similar behaviors in market prices

The first data corresponds to daily evolution of an energy market prices named Y_2. The figure 2 shows the observed data (in blue) and its segmentation (in red). This one contains 21 increasing, decreasing and constant segments. The goal is to detect similar behaviors among this evolution of market prices, then we apply methodology of hierarchical ascending clustering developed in subsection 2.2. We have chosen $\alpha = 0.01$.

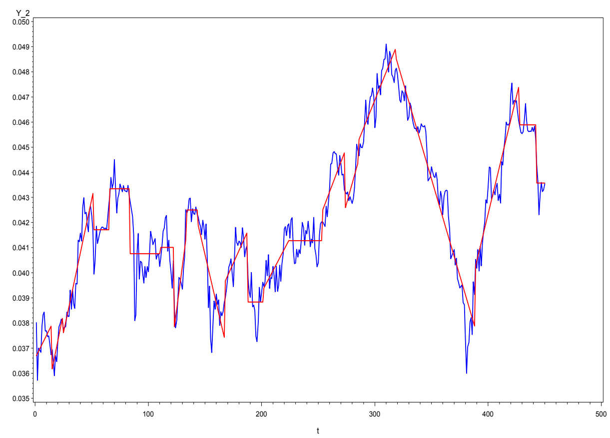

Fig. 2. Evolution of observed market prices and its segmentation

The figure 3 shows the clustering tree. The first both segments aggregated are τ_5 and τ_{21} (in cyan) and their dissimilarity is $d_{(5,21)} = 0.15$ denoted "1-$p$value" on $x-axis$. Then τ_{15} and τ_{16} are mixed at a dissimilarity equal to 0.47 (purple color). We can remark that this cluster is aggregated with the segment τ_{19}.

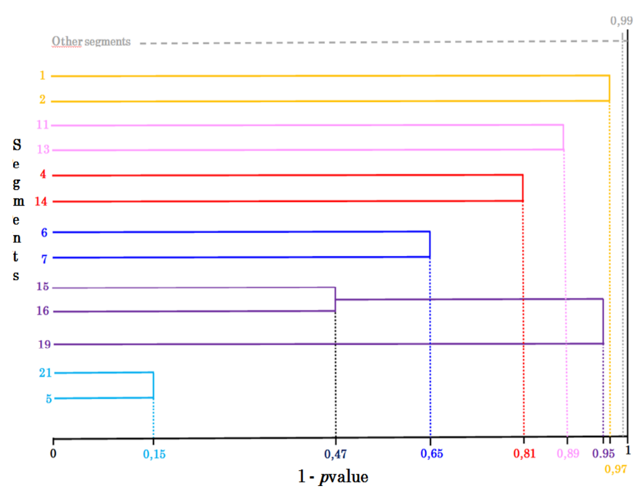

Fig. 3. Clustering tree

The figure 4 shows 14 clusters of 21 segments with $\alpha = 0.01$. For instance, the cluster 4 (red dot) contains two constant segments τ_4 and τ_{14} whereas the cluster 11 (purple cross) has two increasing segments τ_{15} and τ_{16}.

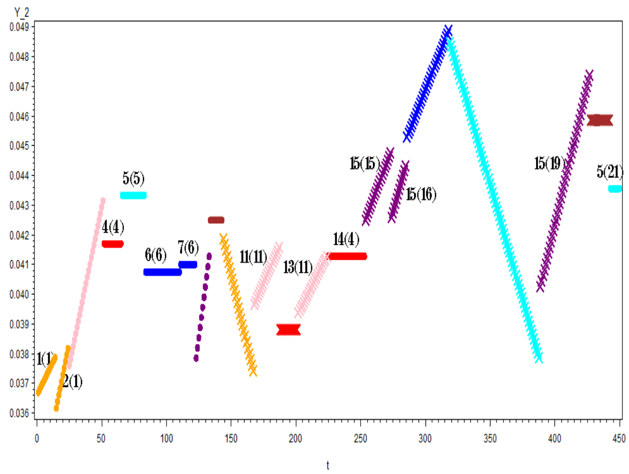

Fig. 4. Clustering of segments with $\alpha = 0.01$

Finally, the figure 5 gives 14 clusters on the original data. The circled segments correspond to clusters with at least two segments, whereas the others contain only one segment.

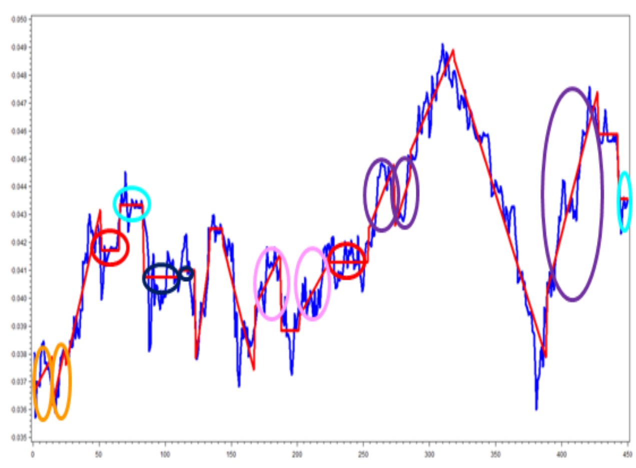

Fig. 5. Evolution of observed market prices and its clustering of segments ($\alpha = 0.01$)

The results obtained of this energy Y_2 provide potential information to analyse behavior market prices, not only for this energy, but also with respect to others energy.

3.2 Detection of abnormal behaviors in marginal costs

The figure 6.a (top-left) shows the hourly evolution of marginal costs on three weeks (in blue) and we remark a daily periodicity. 75 change points have been detected (76 constant segments) and are represented in red on the figure 6.b (top-right). The evolution of observed ordered p-values (in red) and expected p-values (in blue) in figure 6.c (bottom-left) show an intersection with an high level of the order k for the segment number 12. Indeed, $\bar{\pi}(\hat{k}) = 0.96$ with $\alpha = 0.01$. This segment is very atypical. At the opposite, on figure 6.d (bottom-right) the segment number 57 is the most "normal" because $\bar{\pi}(\hat{k}) = 0$.

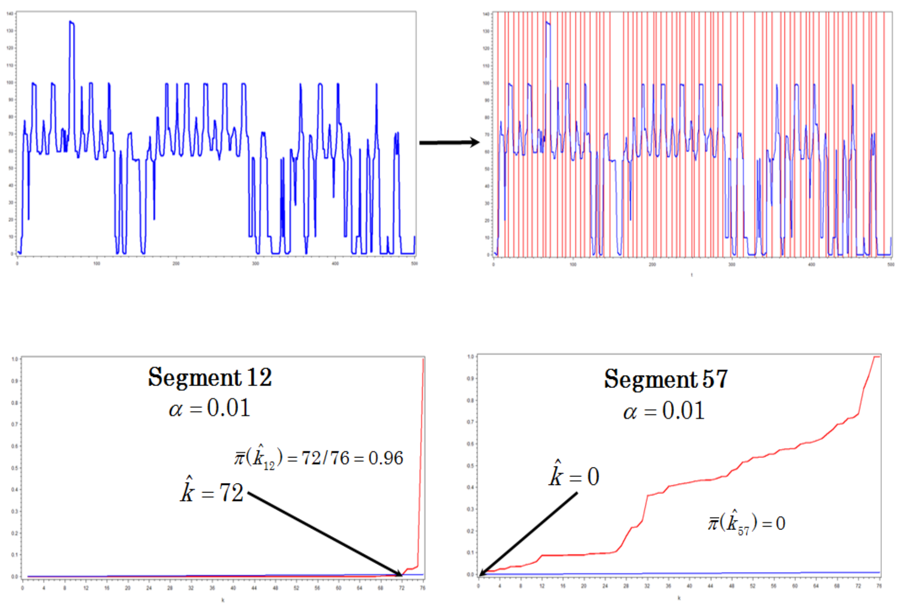

Fig. 6. (a) Marginal cost (b) Change points (c) Atypical seg. (d) No atypical seg.

The figure 7.a (top) shows 76 constant segments of marginal costs (mean of marginal costs). However, it is not possible to detect statistically, if there are atypical segments.

On the figure 7.b (bottom), confidence interval is provided for each segment. For the atypical segment 12, the confidence interval (95%) is very narrow. Notably its lower bound is greater than the upper bound of the all others confidence intervals. At the opposite, the confidence interval of segment 57 is very large and contains all means of others segments. In addition, the segment 22 has the largest confidence interval, whereas the segments 52, 71, 72, 75 and 76 have the more little for a marginal cost of 0.

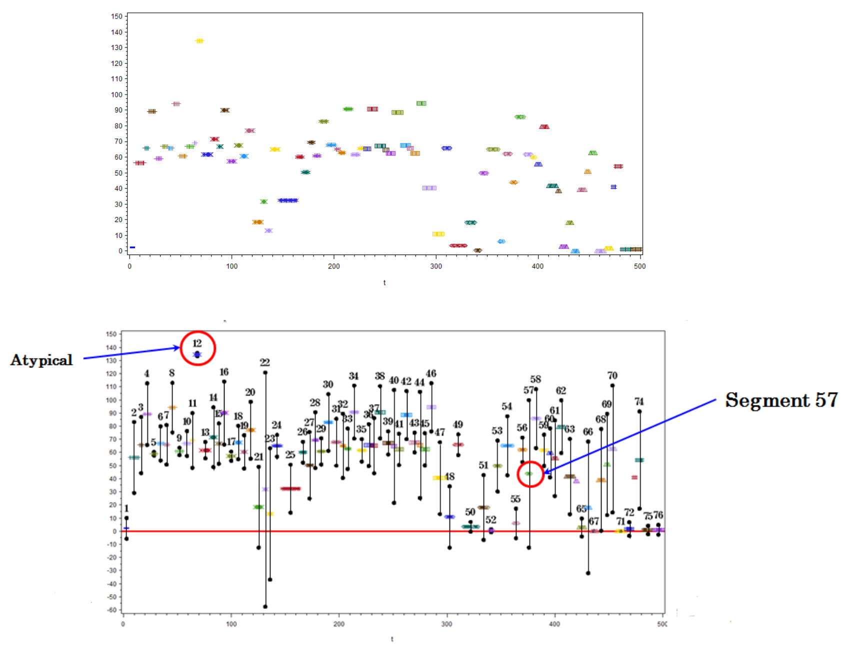

Fig. 7. (a) 76 segments of marginal cost (b) Segments with confidence intervals

4 Contribution, applications and further researches

The two approaches introduced in this paper allow to detect similar behaviors in irregular time series and atypical behaviors in regular time series. These two approaches are applied on segmented time series. The first ones consists in applying a hierarchical ascending clustering to aggregate the segments, in using a controlled stopping rule depending of a threshold α associated to the comparing regression coefficients between two segments. The detection of abnormal behaviors uses the same tests of comparing regression coefficients between segments. For each one, $M-1$ tests are applied and the method of multiple tests is used to measure the proportion of rejected tests. If this one is abnormally high then the segment is atypical. The two applications in section 3, have provided interesting results not only in terms of ability to detect similar or abnormal behaviors in time series, but also it gives a very fruitful information for the experts in energy management. The contribution of two approches is particulary important when there are lot of time series to analyse. Indeed, in frame of simulator in domain of energy management, a task is to detect anomalies of simulated time series. However the number of output time series can be very high. In this case, it is impossible to analyse these data one by one. Then the approach proposed in section 2.3 is very interesting to reduce the time computing of analysis of data. Our further researches will develop an approach to detect similar and atypical behaviors on multivariate time series.

References

1. Arlot, S. & Celisse, A. (2010): Segmentation of the mean of heteroscedastic data via cross-validation, *Statistics and Computing*, pp. 1-20.
2. Benjamini Y. & Hochberg Y. (1995): Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 57, No. 1 (1995), pp. 289-300, Wiley.
3. Derquenne, C. (2011a): An Explanatory Segmentation Method for Time Series, *in Proceedings of Compstat 2010*, Y. Lechevallier & G. Saporta (eds.), 1^{st} Edition, pp. 935-942.
4. Derquenne, C. (2011b): Segmentation of Time Series with Heteroskedastic Components, 58^{th} *World Statistical Congress of ISI*, Dublin, Ireland.
5. Derquenne, C. (2012): Meta-segmentation of time series for searching a better segmentation, *in Proc. of Compstat 2012*, Limassol, Cyprus, pp. 191-204.
6. Derquenne, C. (2013): Clustering of time series via a segmentation approach, IFCS 2013, *Program and Book of Abstracts*, Tilburg, The Netherlands p. 134.
7. Derquenne, C. (2014): Modelling multivariate time series by structural equations modelling and segmentation approach, *in Proc. of Compstat 2014*, Geneva, Switzerland, pp. 459-466.
8. Derquenne, C. (2015): Formation d'un phénomène temporel à l'aide d'un méta-modèle via la segmentation, Actes des 47^{iemes} Journées de Statistique, Lille, France.
9. Guédon, Y. (2008): Exploring the segmentation space for the assessment of multiple change-point models. Institut National de Recherche en Informatique et en Automatique, *Cahier de recherche 6619*.
10. Lavielle, M. and Teyssière, G. (2006): Détection de ruptures multiples dans des séries temporelles multivariées. *Lietuvos Matematikos Rinikinys*, Vol **46**.
11. Rao, CR. and Kleffe, J. (1988): *Estimation of variance components and applications*, North Holland series in statistics and probability, Elsevier.

Valuation of American Put Options with exercise restrictions

Domingos Djinja[1]

Department of Mathematics and Informatics, Faculty of Sciences, Eduardo Mondlane University, Mozambique
(E-mail: celsodjinja@gmail.com)

Abstract. Many authors have been researching American Put Options, particularly on seeking an explicit formula for it or numerical methods with better performances.

In this paper we price American put options with exercise restrictions on weekends. The basic idea relies on removing weekends and shrinking the interval (useful days). Then, jumps may arise on the stock price. Thus, some theory of jump diffusion process is used on pricing these options both analytically and numerically. On this last, an extension to the algorithm presented by B. Kim et al (2013) is presented in order to get the corresponding optimal exercise boundary.
Keywords: American Options, Martingale, Partial differential equation, Finite difference method, Jump diffusion Processes, Early exercise, Optimal Exercise boundary, Critical stock price.

1 Introduction

American options are commonly traded through the world. It is well known that they are a kind of optimal stopping problems since they can be exercised at any time during its lifetime. Moreover, they can be formulated as free boundary value problems. On Peskir & Shiryaev [6] for instance, the conversion from an optimal stopping to a free boundary value problem is explained.

We consider American put options under standard Black-Scholes conditions but with exercise restrictions on weekends. The idea is to remove weekends and shrink the interval. Then, jumps may arise on the stock price. Thus, we use some theory of jump diffusion processes on Cont & Tankov [2] to hedge and price these options. Moreover, in order to get the critical price we extend the algorithm presented by B. Kim et al [5] which relies on solving numerically the Black Scholes equation.

2 The stock price for the problem in analysis

We regard a problem of pricing an American put with exercise restrictions on weekends. We suppose that have a standard brownian motion $[W(t)]$ under a complete probability space $(\Omega, \mathcal{F}, P)$ and $(\mathcal{F}_t)_{t\geq 0}$ is a filtration which satisfy

16^{th} *ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece*

the usual conditions. Furthermore, we consider a financial market with a constant risk-free interest rate $r > 0$ and a stock price $S(t)$ follows a geometric brown motion.

In order to deal with the exception of exercising the American put option during weekends, we suppose the following cases:

1. The stock price is traded continuously at any time;
2. The stock price cannot be traded during the weekends but during the week it is continuously traded.

Case 1: The stock price is traded continuously at any time.

Suppose that we the stock price is a geometric brownian motion, it is continuous and follows the dynamic:

$$dS(t) = \mu S dt + \sigma S dW(t), \qquad t \in [0, T],$$

where μ is the drift, σ is the volatility, $W(t)$ is a standard brownian motion.

Since we are not allowed to exercise the option during the weekends, we will remove the weekends and consider the stock price during the week. Therefore, we may have jumps from Friday to Monday since the price may change during the weekend. As S is a geometric brownian motion, with σ, μ constants, by Björk [1], we have

$$S(t_2) = S(t_{1_-}) \exp\left\{(\mu - \sigma^2/2)(t_2 - t_1) + \sigma[W(t_2) - W(t_1)]\right\},$$

for $t_1 < t_2$ and μ, σ constants.

Thus, if we regard that there are n weekend over the interval $[0, T]$, and we order them as

$$(\tau_1, \tau_2), ..., (\tau_{2n-1}, \tau_{2n}).$$

The jump size at each interval is given by

$$Y_i = exp\{(\mu - \sigma^2/2)(\tau_{2i} - \tau_{2i-1}) + \sigma[W(\tau_{2i}) - W(\tau_{2i-1})]\}.$$

By removing the weekends on the stock price, we will have the following dynamic

$$\frac{dS(t)}{S(t_-)} = \mu dt + \sigma dW(t) + dJ(t), \qquad t \in [0, T] \setminus \{\cup_{i=1}^{n}(\tau_{2i-1}, \tau_{2i})\},$$

where

$$J(t) = \sum_{i=1}^{n(t)}(Y_i - 1), \qquad n(t) \text{ is the number of weekends up to time } t$$

and

$$Y_i = exp\{(\mu - \sigma^2/2)(\tau_{2i} - \tau_{2i-1}) + \sigma[W(\tau_{2i-1}) - W(\tau_{2i-1})]\}.$$

In other to simplify notations and for calculations proposes, We regard that the stock price has the dynamics:

$$\frac{dS(t)}{S(t_-)} = \mu dt + \sigma dW(t) + [Y(t) - 1]dn(t),$$

where $n(t)$ is one if we have jump at time t (i.e. if $t = \tau_{2j}$ the time just after a weekend), it is zero otherwise and $S(t_-)[Y(t)-1]$ represents the jump at time t. We suppose that dW and dn are independent. Therefore,

$$S(t) = S(0) * exp\left[(\mu - \sigma^2/2)t + \sigma W(t) + \sum_{i=1}^{n(t)}(Y_i - 1)\right]. \tag{1}$$

In order to avoid arbitrage, We need the model to be a martingale. Without removing the weekends we have the standard Black-Scholes model which is complete and with a unique martingale measure. By removing the weekends and shrinking the time interval we may have a different scenario, i.e, some jumps on the stock price may appear at known dates. Since we know that $e^{-rt}S(t)$ is a martingale (in this case $\mu = r$ on (1)), it should be natural to it keep so even with the referred possible jumps since the stock price is continuously traded. Then, the jumps must be a martingale. So,

$$E[S(\tau_{i^-})Y_i(t)|\mathcal{F}_s] = S(\tau_{i^-}), \qquad s < t, \quad i = 1, ..., n(T),$$

$\{\mathcal{F}_t\}_{t\geq 0}$ is the information flow up to time t. The last formula is equivalent to

$$E[Y_i(t)|\mathcal{F}_s] = 1. \tag{2}$$

Thus,

$$e^{-rt}E[S(t)|\mathcal{F}_s] = e^{-rt} \cdot S(0) \cdot e^{rt} \cdot e^{\left[\sum_{i=1}^{n(t)}(E[Y_i|\mathcal{F}_s]-1)\right]} = S(0).$$

From condition (2), follows that the interest rate should be zero along the weekend. Therefore,

$$Y_i = exp\{-\sigma^2/2(\tau_{2i} - \tau_{2i-1}) + \sigma[W(\tau_{2i-1}) - W(\tau_{2i-1})]\}.$$

Remark 21 The market is still complete since the stock is traded continuously and the adjustments that we do only are made in order to compute the option price which cannot be exercised during weekends.

Case 2: The stock price cannot be traded during the weekends but during the useful week days it is continuously traded.

As before, we suppose that the stock price is a geometric brownian motion and has the following dynamic:

$$\frac{dS(t)}{S(t_-)} = \mu dt + \sigma dW(t), \qquad t \in [0,T] \setminus \{\cup_{i=1}^{n}(\tau_{2i-1}, \tau_{2i})\},$$

where (τ_{2i-1}, τ_{2i}) $i = 1, 2, ..., n(T)$ are weekends.

Since it is not traded during the weekends, after a weekend we may have a different value from the last one we end up with on the week before. There are many reasons that may be in the origin of this change, for instance, a political decision, a natural catastrophe, a terrorist attack, of course, depending on the asset on trading. Therefore, it is more convenient to consider that there will be a jump from Friday to Monday. Then, we have more random sources than traded assets. So, the market is incomplete and we may have then many martingale measures. However, if we suppose that the jumps are given by a stochastic variable which has log-normal distribution since the stock price has log normal distribution during the useful week days, we would have a bit similar case with the previous one. Thus, the value of the stock price just after a weekend is

$$S(\tau_{2i^+}) = S(\tau_{2i-1^-})e^{a+b\cdot Z(t)},$$

where a, b are constants and $Z(t) \sim N(0, 1)$, i.e, $Z(t)$ has standard normal distribution.

Therefore, a similar argument as on the previous case to avoid arbitrage, we must have the jumps to be martingale

$$E[S(\tau_{2i-1^-})e^{a+b\cdot Z(t)}|\mathcal{F}_s] = S(\tau_{2i-1^-}), \qquad s < t$$

which implies that

$$E[e^{a+bZ(t)}|\mathcal{F}_s] = 1.$$

Consequently, we have $a = -b^2/2$. Since the jumps should reflect the stock price behavior during the weekend if it is traded along this time, then the natural value for b is the corresponding coefficient of a standard brownian motion that we have along the week which is

$$b = [W(t_{i+1}) - W(t_i)] \cdot \sigma, \quad t_i < t_{i+1}.$$

By introducing these possible jumps under martingale measure in $S(t)$, we have

$$S(t) = S(0) \cdot exp\left[(r - \sigma^2/2)t + \sigma W(t) + \sum_{i=1}^{n(t)}(e^{-b^2/2+bZ(\tau_{2i^+})} - 1)\right],$$

where

$$b = \sigma \cdot [W(\tau_{2i}) - W(\tau_{2i-1})], \quad i = 1, 2, .., n(T).$$

Remark 22 Both cases have similar (equal) formulas. However they are different, on the first one the market is complete and the second one not. Nevertheless, by choosing the martingale measure as we did on both cases, we have the same price process. Therefore, from now on we will treat them as a unique case, i.e, the stock price is given by

$$S(t) = S(0) \cdot exp\left[(r - \sigma^2/2)t + \sigma W(t) + \sum_{i=1}^{n(t)}(Y_i - 1)\right].$$

where

$$Y_i = exp\{(-\sigma^2/2)(\tau_{2i} - \tau_{2i-1}) + \sigma[W(\tau_{2i}) - W(\tau_{2i-1})]\}, \quad i = 1, 2, ..., n(T).$$

3 The pricing problem

We will now price an American option when the stock price is given by the last two cases and suppose that the strike price is K. We regard that $P(t, S)$ is the option price. By applying Itô formula for diffusions with jumps on Cont & Tankov [2], we have:

$$dP = \left(\frac{\partial P}{\partial t} + \mu S\frac{\partial P}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 P}{\partial S^2}\right) dt + \sigma S\frac{\partial P}{\partial S}dW(t) + [P(t, S(t)) - P(t, S(t_-))]dn(t).$$

Let us now make a Δ-hedged portfolio and we regard $\delta = \frac{\partial P}{\partial S}$:

$$\Pi(t) = P(t) - \delta \cdot S(t).$$

We have thus,

$$\begin{aligned} d\Pi(t) = dP - \delta dS = \left(\frac{\partial P}{\partial t} + \mu S\frac{\partial P}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 P}{\partial S^2}\right) dt + \sigma S\frac{\partial P}{\partial S}dW(t)+ \\ +[P(t, S(t)) - P(t, S(t_-))]dn(t) - \frac{\partial P}{\partial S}S(\mu dt + \sigma dW(t)+ \\ +(Y - 1)dn(t)) = \\ = \left(\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 P}{\partial S^2}\right) dt + [\Delta P(t, S(t)) - \delta \cdot \Delta S]\, dn(t). \end{aligned}$$

In order to avoid arbitrage, the expected return of the hedged portfolio must be equal to the value of the portfolio invested at risk-free interest rate r. Therefore,

$$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 P}{\partial S^2} + E[\Delta P(t, S) - \delta\Delta S(t)|\mathcal{F}_s] \cdot I_{(\tau_{2i-1},\tau_{2i})}(t) = r(P - \delta S),$$

$s < t$, $I_{(a,b)}(t)$ is the indicator function of the interval (a, b). Since each $Y_i - 1$ is a martingale, we have

$$E[\delta\Delta S(t)|\mathcal{F}_s] = E[\delta S(\tau_{2i-1^-})(Y_i - 1)|\mathcal{F}_s] = \delta S(\tau_{2i-1^-})E[Y_i - 1|\mathcal{F}_s] = 0.$$

Thus, for the $e^{-rt}P(t, S)$ be a martingale, we impose the condition

$$E[\Delta P(t, S(t) \cdot I_{(\tau_{2i-1},\tau_{2i})}|\mathcal{F}_s] = 0, \quad i = 1, 2, ..., n(T),$$

which is equivalent to

$$E[P(\tau_{2i-1^-}, S(\tau_{2i-1^-})Y_i)|\mathcal{F}_s] = P(\tau_{2i-1^-}, S(\tau_{2i-1^-})), \quad i = 1, 2, ..., n(T).$$

Therefore, the pricing problem becomes

$$\begin{cases} \dfrac{\partial P}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 P}{\partial S^2} + rS\dfrac{\partial P}{\partial S} - rP = 0, \quad t \in (0,T) \\ E[P(\tau_{2i-1^-}, SY_i)|\mathcal{F}_s] = P(\tau_{2i-1^-}, S(\tau_{2i-1^-})), \quad i = 1,2,...,n(T), \\ P(T,S) = \max\{K - S,\, 0\}. \end{cases} \tag{3}$$

We will have both analytical and numerical valuation for this problem.

3.1 Analytical valuation

We will now derive a formula for the option price under the stock price defined on the previous section. Consider an American put on the corresponding asset with strike price K and maturity time T. We consider the value of the American put at time $t = T - t'$ as $P_A(t',\ S)$, which is taken on the space $D = \{(t', S) : S \in (0, \infty),\ t' \in [0,T]\}$. There is a critical stock price S^* (exercise boundary) at each time $t \in [0,T]$ such that it is optimal to exercise the option when $S \leq S^*$ and it should continue otherwise. Thus, the American put can be written as

$$P_A(t', S) = \begin{cases} K - S(t), & \text{if } S(t) \leq S^*(t) \\ P_A(t', S) > K - S(t), & \text{otherwise}\ , \end{cases}$$

where $t = T - t'$.

We rewrite the pricing problem as

$$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0, \quad t \in (0,T) \tag{4}$$

satisfying

$$E[P(\tau_{2i-1^-}, SY_i)|\mathcal{F}_s] = P(\tau_{2i-1^-}, S(\tau_{2i-1^-})), \quad i = 1,2,...,n(T), \tag{5}$$

with the terminal and boundary conditions (also known as smooth-pasting conditions)

$$\begin{cases} P(T,S) = \max\{K - S,\, 0\} \\ \lim\limits_{S\to\infty} P(t,S) = 0 \\ \lim\limits_{S\to S^*} P(t,S) = K - S^* \\ \lim\limits_{S\to S^*} \dfrac{\partial P}{\partial S}(t,S) = -1. \end{cases} \tag{6}$$

Following the steps of Gukhal [4], we proved the following theorem:

Theorem 31 *The solution for the problem above is given by*

$$P_A = P_E + Kr\left(\int_0^{\tau_1} e^{-rt}E(I_{\{S\le S^*\}})dt + \sum_{i=1}^{n(T)}\int_{\tau_{2i}}^{\tau_{2i-1}} e^{-rt}E(I_{\{S\le S^*\}})dt\right) -$$
$$-\sum_{i=1}^{n(T)} E[I_{\{S>S^*,YS\le S^*\}}[K - YS(\tau_{2i-1^-}) - P(S(\tau_{2i-1}))]]-$$
$$-\sum_{i=1}^{n(T)} E[I_{\{S\le S^*,\,YS>S^*\}}[P(Y_iS(\tau_{2i-1^-})) - K + S(\tau_{2i-1^-})]],$$

where P_E is the corresponding European option, T is supposed to be a Friday and τ_i correspond to the a beginning or a end of a weekend.

Remark 31 The solution is a sum of an European put, a premium for early exercise and some negative components. These last components correspond to the cases when the stock price jumps from the exercise region to continuation region and vice-versa without across the boundary. We can interpret it as a loss for not exercise the option at that time.

The critical stock price is given when $S(t) = S^*(t)$, thus it is the solution to the following equation

$$P_A(S^*,t) = P_E(S^*,t) + Kr\int_0^t e^{-r\xi}E(I_{\{S(t)\le S^*(t-\xi)\}})d\xi-$$
$$-\sum_{i=1}^{n(t)} E[I_{\{S>S^*,YS\le S^*\}}[K - YS^*(\tau_{2i-1^-}) - P(S^*(\tau_{2i-1}))]]-$$
$$-\sum_{i=1}^{n(t)} E[I_{\{S\le S^*,\,YS>S^*\}}[P(Y_iS^*(\tau_{2i-1^-})) - K + S^*(\tau_{2i-1^-})]].$$

3.2 Numerical valuation

We will rely on the numerical method presented by Kim (2013) et al [5] to price this American put option with jumps. We will use Bermudan options, regarding that it is only exercisable at discrete time points (days) which (τ_{2i-1}, τ_{2i}) $i = 1, 2, ..., n(T)$ correspond to weekends. First of all, we will make some transformations on the expectation of the price variation along the jumps. We have,

$$E[P(\tau_{2i-1^-}, SY_i)|\mathcal{F}_s] = P(\tau_{2i-1^-}, S(\tau_{2i-1^-})), \quad i = 1, 2, ..., n(T),$$

then if we suppose that the price just after the jump is equal to the one just before the jump we will have

$$P(\tau_{2i-1^-}, SY_i) = P(\tau_{2i-1^-}, S(\tau_{2i-1^-})), \quad i = 1, 2, ..., n(T).$$

Therefore, regarding $\beta(t,T)$ the critical stock price, we have the following free boundary problem:

$$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0, \quad t \in (0,T),\ S \in (\beta,\, \infty)$$

satisfying

$$P(\tau_{2i-1^-}, SY_i) = P(\tau_{2i-1^-}, S(\tau_{2i-1^-})), \quad i = 1, 2, ..., n(T),$$

moreover

$$P(T, S, T) = \max\{K - S,\, 0\}, \qquad \beta(T,T) = K,$$

and the boundary conditions

$$\lim_{S\uparrow\infty} P(t,S,T) = 0,$$
$$\lim_{S\uparrow\beta} P(t,S,T) = K - \beta(t,T),$$
$$\lim_{S\downarrow\beta} \frac{\partial P(t,S,T)}{\partial S} = -1.$$

For a numerical treatment, we need to generate the jumps. Since they contain brownian motion or normal distribution in their formula and a deterministic part, we only need to generate brownian motion paths. We chop the interval $[0,T]$ in N equal subintervals such that $t_n = n\Delta t$, $n = 0, 1, ..., N$, $\Delta t = \frac{T}{N}$. We interpret each t_n as one day. By using the random walk construction in P. Glasserman (chapter 3.2, [?]), we have

$$W(t_{i+1}) = W(t_i) + \sqrt{t_{i+1} - t_i}Z_{i+1}, \qquad i = 0, ..., N-1,$$

where $0 = t_0 < t_1 < ... < t_n$, $W(0) = 0$, Z_1, $Z_2, ..., Z_n$ are independent normal distribution. Thus,

$$W(t_{i+1}) - W(t_i) = \sqrt{t_{i+1} - t_i}Z_{i+1}, \qquad i = 0, ..., N-1.$$

We have therefore,

$$\begin{aligned} Y_i &= exp[-\sigma^2/2(\tau_{2i} - \tau_{2i-1}) + \sigma\sqrt{\tau_{2i} - \tau_{2i-1}}Z_i] \\ &= exp\left[\frac{3T}{N}(-\sigma^2/2) + \sigma\sqrt{\frac{3T}{N}} \cdot Z_i\right], \qquad i = 1, 2, ..., N, \end{aligned}$$

where Z_1, Z_2, ..., Z_n are independent standard normal random distribution. At each weekend, we can simulate a sample of jumps and determine its mean. Then, we regard the sample mean as a jump at each weekend.

We follow then the algorithm for pricing a standard American put option (in Kim 2013 cited above), but with some adjustments because of the possible jumps on the stock price. Let Ω_e, Ω_c be the exercise and continuous region, respectively, (S_{min}, S_{max}) the interval of the stock price. By setting $Q = \sqrt{P - K + S}$, we see that $Q = 0$ on the free boundary and Ω_e.

The method consists on solving the backward time Black-Scholes equation. We consider a variable weekdayforT, the starting useful day T as the number of days backward up to Sunday. We chop the interval of the stock price, $[S_{min}, S_{max}]$ in M equal subintervals, such that $S_i^n = S_i(t_n)$ $i = 0, 1, 2, ...M$, and $\Delta S = \frac{S_{max}-S_{min}}{M}$,

$$S_i^n = \beta_n + \rho\Delta S + i\Delta S, \qquad n = N,\ N-1, ..., 0, \quad i = 0, 1, ..., M,$$
$$0 < \rho < 1,\ \text{(constant)}$$

and the option price $P(t_n, S_i^n, T) = P_i^n$. S_i^n is computed after the free boundary β_{n+1} be computed, and $\beta_N = K$. Then we proceed by following the steps:

1. Determine the option price P_i^{n-1}, $i = 1, 2, ..., M$ explicitly from P_i^n. The price P_i^n of the American put option is a discrete solution to the discrete Black Scholes equation:

$$\frac{P_i^n - P_i^{n-1}}{\Delta t} + \frac{1}{2}\sigma^2 S_i^2 \frac{P_{i+1}^n - 2P_i^n + P_{i-1}^n}{\Delta S^2} + rS_i \frac{P_{i+1}^n - P_i^n}{\Delta S} - rP_i^n = 0$$

$$\text{for } n = N,\ N-1,\ ..., 1, \qquad S_i = S_i^n, \quad i = 1, 2, ..., M-1,$$
$$P_i^N = 0 \qquad \text{for} \quad i = 0, 1, 2, ..., M.$$

2. Find P_0^{n-1} by solving:

$$\frac{P_0^n - P_0^{n-1}}{\Delta t} + \frac{1}{2}\sigma^2 S_0^2 \left(\frac{\frac{P_1^n - P_0^n}{\Delta S} - \frac{P_0^n - P_{-1}^n}{\rho \Delta S}}{\frac{\Delta S + \rho\Delta S}{2}} \right) + rS_0 \left(\frac{P_1^n - P_{-1}^n}{\Delta S + \rho\Delta S} \right) - rP_0^n = 0$$

$$P_{-1}^n = K - \beta_n.$$

Assuming the computational domain is large, we impose zero boundary condition
$P_M^n = 0, \quad n = N,\ N-1, ..., 0.$

3. Determine $\beta_{n-1} = \dfrac{S_0^{n-1}}{1+\xi}$, where ξ is the solution for the equation

$$\xi^3 - \left\{ \ln \frac{S_0^{n-1}}{\beta_n} + (\sigma^2 + r)\Delta t \right\} \xi^2 + 3\sigma^2 \Delta t \xi - \frac{3\sigma^2 \Delta t}{\sqrt{rK}} Q(t_{n-1}, S_0^{n-1}, T) = 0, \tag{7}$$

which has a unique real root $\xi \in (0, 1)$.

4. Change the values of S^{n-1} from old $S_i^{n-1} = \beta_n + \rho\Delta S + i\Delta S$ to new $S_i^{n-1} = \beta_{n-1} + \rho\Delta S + i\Delta S$ for $i = 0, 1, 2, ..., M$. If we archive a weekend day (Sunday), then we determine the jump (we simulate n jumps and take expectation), we ignore Sunday and Saturday. Then, we determine the value on the previous useful day, Friday by

$$S_j^{n-3} = S_j^n / Y_n, \qquad j = 0, ..., M,$$

and we set the price on Friday by

$$P_j^{n-3} = P_j^n$$

for $j = 0, ...M-1$. Then, we determine the value of $\beta[n-3]$ from $\beta[n]$ by solving the equation (7) and by setting

$$\beta_{n-3} = \frac{S_0^{n-3}}{1+\xi}.$$

If we finish step 4, then we repeat the running step 1 through 4 until t_0.

A simulation with the following data: $T = 0.5,\ \sigma = 0.2,\ r = 0.05,\ K = 1,\ \rho = 0.4,\ MaxS = 30$ give the plots in Fig. 1:

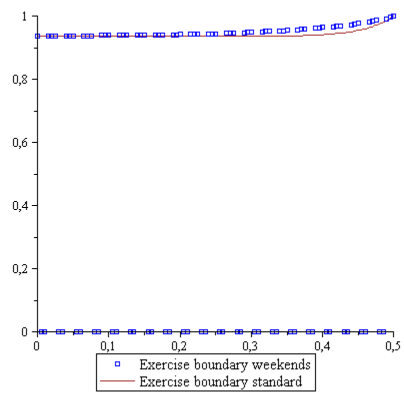

Fig. 1 Optimal exercise boundary for the standard case and the one with exercise restrictions on weekends

Remark 32 We have a continuous line during the useful week days and we set 0 along the weekends. When there is no jump, the line looks continuous. The jumps are random but they may arrive only on weekends (where we set the critical stock price as zero).

4 Conclusion

We found a similar result compared to the general cases of jump diffusion models. The critical stock price is peace-wise continuous and have similar behavior during useful week compared to the standard case. However, it is bigger than the standard case. Thus, there are more chances to exercise earlier the option on the case with restriction on weekends than the standard one.

References

1. Björk, T. *Arbitrage Theory in Continuous Time.* 3rd edition, Oxford, 2009.

2. CONT, R. & TANKOV, P. *Financial modeling with jump processes.* Chapman & HallICRC financial mathematics series, 2004.
3. GLASSERMAN, P. *Monte Carlo Methods in Financial Engineering.* Springer, New York, USA, 2003
4. GUKHAL, C. R. Analytical valuation of American options on jump-diffusion processes. *Mathematical Finance, Vol. 11, No. 1 (January 2001)*, pp 97-115.
5. KIM, B. J., AHN, C. & CHOE, H. J. Direct computation for American put option and free boundary using finite difference method. *Japan J. Indust. Appl. Math.* pp 21-37, 2013.
6. PESKIR, G. & SHIRYAEV, A. *Optimal stopping and free-boundary problems.* Resuahkrib, Zurich, 2006.

Ackonledgements

Special thanks to Professor Erik Ekstrom who contributed with useful comments.

Estimating Functions in the Presence of a Nuisance Parameter

Maria Douli

Department of Statistics and Insurance Sciences, University of Piraeus, 80 Karaoli & Dimitriou str., 185 34 Piraeus, Greece
(E-mail: madouli@hotmail.com)

Abstract. In this work lower bounds for the estimation of a real-valued parameter α of parametric models in the presence of a nuisance parameter β are presented. In particular, an optimal estimating function which provides an estimator for α with the smallest possible variance was developed, when for the elimination of the nuisance parameter an estimator which is independent of the sample is available. The resulting estimating function is better as compared with Stein's adaptive score function.
Keywords: Estimating functions; nuisance parameters; adaptive score function.

1 Introduction

Let $P_{\underline{\theta}}, \underline{\theta} \in \Theta$ be a family of probability measure on a measurable space $(\mathcal{X}, \mathcal{A})$, where Θ is an open subset of $\mathbb{R}^2$. Let also $f(x;\underline{\theta}) = \frac{dP_{\theta}}{d\mu}(x)$ be the probability density function of $P_{\underline{\theta}}$ in terms of a σ-finite measure μ. We write $\underline{\theta} = (\alpha, \beta)^t$ where $\alpha, \beta \in \mathbb{R}$. The aim is to estimate the structural parameter α while the second part β is considered as nuisance parameter. Let us use the notation

$$\underline{J}(x;\underline{\theta}) = (J_1(x;\underline{\theta}), J_2(x;\underline{\theta}))^t = (\frac{\partial \log f(x;\underline{\theta})}{\partial \alpha}, \frac{\partial \log f(x;\underline{\theta})}{\partial \beta})^t$$

where the corresponding information matrix is

$$I(\underline{\theta}) = E(\underline{J}(x;\underline{\theta})\underline{J}^t(x;\underline{\theta})) = \begin{bmatrix} I_{11}(\underline{\theta}) & I_{12}(\underline{\theta}) \\ I_{21}(\underline{\theta}) & I_{22}(\underline{\theta}) \end{bmatrix}.$$

The existence of optimal estimating function(O.E.F.) for α in the presence of nuisance parameter β has been studied by many authors. Stein[10] introduced the concept of efficient score function

$$\hat{J}_1(x;\underline{\theta}) = J_1(x;\underline{\theta}) - I_{12}(\underline{\theta}) I_{22}^{-1}(\underline{\theta}) J_2(x;\underline{\theta}). \quad (1)$$

Neyman[9], Bickel *et al.*[2] proved that the use of a $\sqrt{n}$ consistent estimator $\hat{\beta}_n$ for β results to an asymptotically efficient estimator for α with asymptotic variance

16^{th} *ASMDA Conference Proceedings, 30 June – 4 July 2015, Piraeus, Greece*

$$\hat{I}(\underline{\theta}) = I_{11}(\underline{\theta}) - I_{12}(\underline{\theta}) I_{22}^{-1}(\underline{\theta}) I_{12}(\underline{\theta}). \qquad (2)$$

Lindsay[8], Liang[7] defined as Fisher information for α in the presence of β the quantity

$$I(\alpha;\beta) = \mathrm{Inf}_{c\in\mathbb{R}} E[\frac{\partial \log f}{\partial \alpha} - c\frac{\partial \log f}{\partial \beta}]^2. \qquad (3)$$

Bhapkar[1] generalized (2) for the multivariate case. The optimality is achieved for $c = I_{12}I_{22}^{-1}$. Thus the optimal estimating function for the estimation of α in the presence of β is Stein's efficient score (1). Let us assume that an estimator $\hat{\beta}_m$ for β from another sample $\underline{y} = (y_1, y_2, ..., y_m)$ is available where the total sample of size $n+m$ consists of dependent observations.
We emphasize that we do not demand the knowledge of the sample $\underline{y}$ but only of the estimator $\tilde{\beta}_m$. The problem we face is how to use the estimator $\tilde{\beta}$ in the estimation of α. Provided that the sample $\underline{y}$ is not available, we can not write the likelihood function in order to proceed in the usual way. With the use of the sample $\underline{x}$ we can have the maximum likelihood estimator $\hat{\beta}_n$ of β. the aim of the paper is to find the O.E.F. for the estimation of α when for the elimination of β we shall use a pooled type estimator

$$\hat{\beta}_{n+m} = w_1\hat{\beta}_n + w_2\tilde{\beta}_m \qquad (4)$$

of the following structure

(i) $w_1 + w_2 = 1$ where $0 < w_2 \leq 1$.The case $w_2 = 0$ correspond to the case where we estimate α and β from the same sample and the O.E.F. is (1).
(ii)The estimator $\hat{\beta}_n$ is the maximum likelihood estimator for β from the sample $\underline{x}$.
(iii) The estimator $\tilde{\beta}_m$ is a consistent and asymptotically normal estimator for β based on the sample $\underline{y}$ with asymptotic variance v.
(iv) Concerning the sample sizes we suppose that they converge to infinity in such a way that

$$\lim_{n,m\to\infty}(n/m) = r, \ where \ 0 \leq r \leq \infty.$$

The term optimal estimating function understood as the estimating function (E.F.) which provides as a root a consistent and asymptotically normal distribution with the smallest possible asymptotic variance. The search of the O.E.F. will take place within the set $\mathcal{G}_{\underline{\theta}}$ which includes the estimating functions $g(x;\underline{\theta})$ which satisfy the following regularity properties (Godambe [3],[4])
R1.$Eg(x;\underline{\theta}) = 0$, for all $\underline{\theta} \in \Theta$.
R2. $Eg^2(x;\underline{\theta}) < \infty$, for all $\underline{\theta} \in \Theta$.
R3. $g(x;\underline{\theta})$ is continuously differentiable for $\underline{\theta} \in \Theta$.
R4.$\int g(x;\underline{\theta})f(x,\underline{\theta})dx$ is differential with respect to $\underline{\theta}$ under the integral sign.
R5. $E_{\underline{\theta}}(\frac{\partial g}{\partial \alpha}) \neq 0$ and $E_{\underline{\theta}}(\frac{\partial g}{\partial \beta}) \neq 0$ for all $\underline{\theta} \in \Theta$.

R6. There are $M_i(x;\underline{\theta}),\ i = 1,2,3$ with $E_{\underline{\theta}}M_i(x;\underline{\theta}) < \infty$ such that in a neighborhood $B(\underline{\theta}, r)$ of the true value $\underline{\theta}$ holds that

$$|\frac{\partial^2 g}{\partial \alpha^2}| < M_1(x;\underline{\theta}),\ |\frac{\partial^2 g}{\partial \beta^2}| < M_2(x;\underline{\theta}),\ |\frac{\partial^2 g}{\partial \alpha \partial \beta}| < M_3(x;\underline{\theta}).$$

For such a $g \in \mathcal{G}_{\underline{\theta}}$ we consider the equation

$$\sum_{i=1}^{n} g(x_i;\alpha,\hat{\beta}_{n+m}) = 0 \qquad (5)$$

which will give as a solution and estimator for α.

In terms of the notation we denote the maximum likelihood estimator of $\underline{\theta} = (\alpha,\beta)$ as $\hat{\underline{\theta}}_n = (\hat{\alpha}_n,\hat{\beta}_n)$. Also D stands for the determinant of the information matrix $I(\underline{\theta})$ and $I_{2\times 2}$ for the 2×2 identity matrix. We also define the following column matrices

$$\underline{C}(\underline{\theta}) = (C_1(\underline{\theta}),C_2(\underline{\theta}))^t = \lim_{n\to\infty} \text{cov}\big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g(x_i;\underline{\theta}),\sqrt{n}(\hat{\underline{\theta}}_n - \underline{\theta})\big) \text{and } R(\underline{\theta}) = \begin{bmatrix}\text{cov}(J_1,g)\\ \text{cov}(J_2,g)\end{bmatrix}.$$

The symbol $\Box$ denotes the end of a proof.

The rest of the paper is organized in the following way. In Section 2 the main results are developed, in Section 3 the results are discussed. Finnaly in Section 4 we apply the results for the Gamma distribution.

2 Main Results

2.1 Consistency

Theorem 1. Let $g(x;\underline{\theta})$ be an estimating function from $\mathcal{G}_{\underline{\theta}}$ which satisfies the following conditions
(1) $E_{\underline{\theta}}g(\text{x};\theta') = 0$ only when $\underline{\theta} = \underline{\theta}'$
(2) the E.F. g is a monotone function for α.
If $\tilde{\beta}_m$ is a consistent solution for β then the equation $\sum_{i=1}^{n} g(x_i;\alpha,\tilde{\beta}_m) = 0$ possesses a consistent solution for α.
Proof.

Let (α_o,β_o) be the true value and $\epsilon > 0$. Let also assume that the $g(x;\alpha,\beta)$ is decreasing for α. Since $E_{(\alpha_o,\beta_o)}g(X;\alpha_o,\beta_o) = 0$ from the monotonicy of g we have

$$\frac{1}{n}\sum_{i=1}^{n} g(x_i;\alpha_o+\epsilon,\tilde{\beta}_m) \to E_{(\alpha_o,\beta_o)}g(X;\alpha_o+\epsilon,\beta_o) = c < 0 \text{ as } n,m\to\infty.$$

which means that

$$P[|\frac{1}{n}\sum_{i=1}^{n} g(x_i;\alpha_o+\epsilon,\tilde{\beta}_m) - c| < \epsilon] \to 1 \text{ as } n,m\to\infty$$

For $\epsilon = -c$ we conclude that

$$P[-\epsilon < \frac{1}{n}\sum_{i=1}^{n} g(x_i;\alpha_o+\epsilon,\tilde{\beta}_m) < 0] \to 1 \text{ as } n,m\to\infty$$

or

$$P[\frac{1}{n}\sum_{i=1}^{n} g(x_i;\alpha_o+\epsilon,\tilde{\beta}_m)<0]\to 1 \text{ as } n,m\to\infty. \quad (6)$$

In a similar way from the relation

$$\frac{1}{n}\sum_{i=1}^{n} g(x_i;\alpha_o-\epsilon,\tilde{\beta}_m)\to E_{(\alpha_o,\beta_o)}g(X;\alpha_o-\epsilon,\beta)=c<0 \text{ as } n,m\to\infty$$

we obtain

$$P[\frac{1}{n}\sum_{i=1}^{n} g(x_i;\alpha_o-\epsilon,\tilde{\beta}_m)>0]\to 1 \text{ as } n,m\to\infty. \quad (7)$$

Let us define the sets

$$S_{n,m}=\left\{(\underline{x}_n,\underline{y}_m):\frac{1}{n}\sum_{i=1}^{n} g(\mathrm{x_i};\alpha_\mathrm{o}+\epsilon,\tilde{\beta}_\mathrm{m})<0,\frac{1}{n}\sum_{i=1}^{n} f(\mathrm{x_i};\alpha_\mathrm{o}-\epsilon,\tilde{\beta}_\mathrm{m})>0\right\}$$

and

$$A_{n,m}=\left\{(\underline{x}_n,\underline{y}_m);\text{a solution } \hat{\alpha}_n \text{ exists of } :\frac{1}{n}\sum_{i=1}^{n} g(\mathrm{x_i};\alpha,\tilde{\beta}_\mathrm{m})=0,\alpha_o-\epsilon<\hat{\alpha}_n<\alpha_o+\epsilon\right\}$$

From (6) and (7) we have that

$$P(S_{n,m})\to 1. \quad (8)$$

Since $S_{n,m} \subset A_{n,m}$ from (8) we have that

$$P(A_{n,m})\to 1. \quad (9)$$

Thus

$$P_{(\alpha_o,\beta_o)}(\alpha_o-\epsilon<\hat{\alpha}_n<\alpha_o+\epsilon)\to 1.$$

The proof is similar when we asume that $g(x;\alpha,\beta)$ is increasing for α. □

Lemma 2.1 *It holds that*

$$C(\underline{\theta})=I^{-1}(\underline{\theta})R(\underline{\theta}). \quad (10)$$

Proof. Expanding in Taylor series we obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\underline{J}(x_i;\underline{\theta})=-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\underline{J}(x_i;\underline{\theta})}{\partial\underline{\theta}^t}\sqrt{n}(\hat{\underline{\theta}}_n-\underline{\theta})+o_p(1).$$

Applying the central limit theorem at the expression

$$\frac{1}{\sqrt{n}}\left[\sum_{i=1}^{n}\underline{J}(x_i;\underline{\theta})\ \sum_{i=1}^{n} g(x_i;\underline{\theta})\right]=\begin{bmatrix}-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\underline{J}(x_i;\underline{\theta})}{\partial\hat{\underline{\theta}}} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\sqrt{n}(\hat{\underline{\theta}}_n-\underline{\theta})\\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n} g(x_i;\underline{\theta})\end{bmatrix}+o_p(1)$$

and equating the variance covariance matrices we obtain (10). □

Corollary 2.1 *It holds that*

$$\mathrm{cov}(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\underline{J}(x_i;\underline{\theta}), \sqrt{n}(\underline{\hat{\theta}}_n - \underline{\theta})^t) = I_{2\times 2}.$$

For $g \in \mathcal{G}_{\underline{\theta}}$ we call $\tilde{g}$ the orthogonal projection of g into the space $\mathcal{H} = \mathrm{span}\,\{J_1, J_2\}$ and $\tilde{g}^{\perp}$ the orthogonal complement of g. We call $V(g)$ the asymptotic variance of the consistent estimator $\hat{\alpha}_n$ which is obtained as a solution of the equation (5).

Theorem 2.1 Let $g \in \mathcal{G}_{\underline{\theta}}$. It holds that

$$\mathrm{V}(g) > \mathrm{V}(\tilde{g}). \tag{11}$$

Proof.

We write $g = \tilde{g} + \tilde{g}^{\perp}$ with $E_{\theta} g = E_{\theta}\hat{g} + E_{\theta}\hat{g}^{\perp}$. The following holds

$$\mathrm{var}(g) = \mathrm{var}(\tilde{g}) + \mathrm{var}(\tilde{g}^{\perp}) \geq \mathrm{var}(\tilde{g}). \tag{12}$$

Also

$$E_{\underline{\theta}}\frac{\partial g}{\partial \alpha} = -E_{\underline{\theta}}(gJ_1) = -E_{\underline{\theta}}(\tilde{g}J_1) = E_{\underline{\theta}}\frac{\partial \tilde{g}}{\partial \alpha}. \tag{13}$$

Similarly

$$E_{\underline{\theta}}\frac{\partial g}{\partial \beta} = E_{\underline{\theta}}\frac{\partial \tilde{g}}{\partial \beta} \tag{14}$$

and obviously

$$\mathrm{cov}(J_i, g) = \mathrm{cov}(J_i, \tilde{g}), i = 1, 2. \tag{15}$$

Let now $\hat{a}_n$ be a consistent solution of the equation

$$\sum_{i=1}^{n} g(x_i; \hat{\alpha}_n, \tilde{\beta}_{n+m}) = 0 \tag{16}$$

Next we expand the left-hand side of (5) around the true value (α, β) and we solve for $\sqrt{n}(\hat{\alpha}_n - \alpha)$. We yield the expression

$$\sqrt{n}(\hat{\alpha}_n - \alpha) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g(x_i;\alpha,\beta) + \frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(x_i;\alpha,\beta)}{\partial \beta}\sqrt{n}(\tilde{\beta} - \beta)}{-[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial \alpha}g(x_i;\alpha,\beta)]} + o_p(1).$$

For the algebraic manipulations see Lehmann[6]. With the use of central limit theorem and Slutcky's theorem we see that $\sqrt{n}(\hat{\alpha}_n - \alpha)$ converges in distribution to

$$N(0, \mathrm{V}(g)) \text{ as } n \to \infty$$

where

$$\mathrm{V}(g) = \frac{E(g^2) + 2CEg'_\beta + (w_1^2 \mathrm{v}_1 + r w_2^2 \mathrm{v})(Eg'_\beta)^2}{(Eg'_\alpha)^2} \tag{17}$$

where v_1 is the asymptotic distribution of $\hat{\beta}_n$. From expressions (12) to (15) we conclude (11). □

From the last theorem we conclude that it is enough to focus on estimating functions of the form

$$\lambda_1 J_1 + \lambda_2 J_2 \ \text{ for } \ \lambda_1, \lambda_2 \in \mathbb{R}.$$

From the expression (17) we can see that for any constant a we have $\mathrm{V}(ag) = \mathrm{V}(g)$. Thus it is enough to focus on estimating functions of the form

$$g_\lambda = J_1 + \lambda J_2. \tag{18}$$

Theorem 2.2 *The optimal estimating function in $\mathcal{G}_{\underline{\theta}}$ is obtained for*

$$\hat{\lambda} = -\frac{I_{12}}{I_{22} + \frac{1}{w_2 r \mathrm{v} - w_1 \frac{I_{11}}{D}}}. \tag{19}$$

Proof. For g_λ as in (18) and with the use of Corollary 2.1, the resulting asymptotic variance becomes

$$\mathrm{V}(\lambda) = \frac{(I_{11} + 2I_{12}\lambda + I_{22}\lambda^2) + \Omega(I_{12} + I_{22}\lambda)^2 - 2w_1(I_{12} + I_{22}\lambda)\lambda}{(I_{11} + I_{12}\lambda)^2}$$

or

$$\mathrm{V}(\lambda) = \frac{(I_{11} + \Omega I_{12}^2) + 2I_{12}(w_2 + \Omega I_{22})\lambda + I_{22}(w_2 - w_1 + I_{22}\Omega)\lambda^2}{(I_{11} + I_{12}\lambda)^2} \tag{20}$$

where

$$\Omega = w_1^2 \frac{I_{11}}{D} + w_2^2 r \mathrm{v}. \tag{21}$$

Differentiating (20) for λ we obtain

$$\hat{\lambda} = -\frac{I_{12}(\Omega D - w_1 I_{11})}{I_{22}(\Omega D - w_1 I_{11}) + w_2 D}.$$

Finally replacing Ω from (21), $\hat{\lambda}$ can be written as in (19).

In order prove that the function $V(\lambda)$ in (20) takes on a minimum value for $\hat{\lambda}$ given in (19), instead of working with the second derivative and with tedious calculations we can argue on the following way: The function $V(\lambda)$ has a vertical asymptote (increases to infinite because $V(\lambda)$ is always positive) at the point $\lambda = -I_{11} I_{12}^{-1}$ and it has as an horizontal asymptote because both enumerator and denominator are a second degree polynomial. Thus $\hat{\lambda}$ is point of minimum because if it were a local maximum, between the local maximum and the asymptote it should be a local minimum. But the $V'(\lambda) = 0$ has only one solution. □

From the last theorem we conclude that the optimal estimating function is

$$J_{\text{opt}} = J_1 - \frac{I_{12}}{I_{22} + \frac{1}{w_2 r\text{v} - w_1 \frac{I_{11}}{D}}} J_2. \tag{22}$$

Theorem 2.3 *The optimal asymptotic variance is*

$$\text{V}_{opt} = (I_{11} - \frac{I_{12}^2}{I_{22} + \frac{1}{r\text{v}}})^{-1}. \tag{23}$$

Proof. We write the enumerator of (20) as

$$E = I_{11} + \Omega I_{12}^2 + Q_1 + Q_2 \tag{24}$$

where

$$Q_1 = 2I_{12}(w_2 + \Omega I_{22})\lambda \tag{25}$$

and

$$Q_2 = I_{22}(w_2 - w_1 + I_{22}\Omega)\lambda^2. \tag{26}$$

Replacing $\hat{\lambda}$ in (25) and (26) we obtain

$$Q_1 = -2I_{12}^2\Big[\Omega - \frac{w_1 w_2 I_{11}}{I_{22}(\Omega D - w_1 I_{11}) + w_2 D}\Big] \tag{27}$$

and

$$\begin{aligned} Q_2 = I_{12}^2\Big[\Omega + \frac{(w_2 - w_1)(\Omega D - w_1 I_{11}) - 2w_2\Omega D}{I_{22}(\Omega D - w_1 I_{11}) + w_2 D} \\ + w_1 w_2 D \frac{(w_2 - w_1)I_{11} + D\Omega}{[I_{22}(\Omega D - w_1 I_{11}) + w_2 D]^2}\Big]. \end{aligned} \tag{28}$$

With the use of (27) and (28) the enumerator E becomes

$$E = I_{11} - \frac{I_{12}^2(\Omega D - w_1 I_{11})}{I_{22}(\Omega D - w_1 I_{11}) + w_2 D} + I_{12}^2 w_1 w_2 D \frac{(w_2 - w_1)I_{11} + D\Omega}{[I_{22}(\Omega D - w_1 I_{11}) + w_2 D]^2}. \tag{29}$$

Replacing (29) in (20) and after some calculations we obtain

$$\text{V}_{\text{opt}} = \frac{I_{22}(\Omega D - w_1 I_{11}) + w_2 D + I_{12}^2 w_1 w_2}{[(w_2 - w_1)I_{11} + \Omega D]D}.$$

Replacing Ω from (21) finally we obtain

$$\text{V}_{\text{opt}} = \frac{r\text{v} I_{22} + 1}{r\text{v} D + I_{11}}.$$

The last relation implies (23). □

3 Discussion

The J_{opt} is the analogue to Stein's efficient score function when a pooled type estimator of the form (4) is used for the elimination of β. In fact, since J_{opt} depends on w_2 it consists a class of estimating functions. However all estimating functions (22) result to the same variance (23) which is independent of w_2. The simplest in structure is the one for $w_2 = 1$ which corresponds to the case where for the elimination of β we are based only on the estimator $\tilde{\beta}_m$. In this case the optimal estimating function becomes

$$J^o_{\mathrm{opt}} = J_1 - \frac{I_{12}}{I_{22} + \frac{1}{r\mathrm{v}}} J_2. \tag{30}$$

This is a very interesting result which claims that the optimal estimating function J^o_{opt} which is obtained when we eliminate the nuisance parameter β with the help of an independent sample it can not be improved if for the elimination of β we utilize the available sample $\underline{X}$. The J^o_{opt} does not share the same properties with Stein's score $\hat{J}_1$. Indeed we can see that

$$\mathrm{V}^{-1}_{\mathrm{opt}} < \mathrm{var}(J^o_{\mathrm{opt}}).$$

The J^o_{opt} is not orthogonal to J_2. Indeed

$$\mathrm{cov}(J^o_{opt}, J_2) = \frac{I_{12}}{1 + r\mathrm{v}I_{22}}.$$

Also

$$E\frac{\partial J^o_{\mathrm{opt}}}{\partial \alpha} \neq \mathrm{var}(J^o_{\mathrm{opt}}).$$

When the second sample increases faster compared to the first sample such that $\frac{n}{m} = o(1)$, i.e. $r = 0$, then from (30) we observe that J^o_{opt} becomes the J_1 score function with variance I^{-1}_{11}. This case corresponds to the case where β is known. When the first sample converges faster compared to the second sample such that $\frac{m}{n} = o(1)$, i.e. $r = \infty$ then J^o_{opt} becomes Stein's estimating function with corresponding variance

$$[I^{-1}_{11} - I^2_{12}I^{-1}_{22}]^{-1}.$$

For $0 < r < \infty$ we have

$$I^{-1}_{11} < \mathrm{V}_{\mathrm{opt}} < [I^{-1}_{11} - I^2_{12}I^{-1}_{22}]^{-1}.$$

4 Application

Let us assume that a sample $\underline{X}$ from the Gamma(α, β) distribution with pdf

$$f(\mathrm{x}; \alpha, \beta) = \frac{x^{\alpha-1}\exp(-\frac{x}{\beta})}{\Gamma(\alpha)\beta^{\alpha}}$$

is available. The aim is to estimate the parameter α while β is considered as a nuisance parameter. Furthermore we assume that the method of moments estimator $\tilde{\beta}_n$ of β is available from a sample $\underline{Y}$ independent of the sample $\underline{X}$. For simplicity purposes we assume $r = 1$. Utilizing the asymptotic variance-covariance matrix of $\sqrt{n}(\hat{\alpha}_n, \hat{\beta}_n)$ where $\hat{\alpha}_n$ and $\hat{\beta}_n$ are the MLEs of α and β respectively given by [5], we can find that Stein's optimal asymptotic variance V_1 when we estimate both α and β from the same sample $\underline{X}$ is given by

$$V_1(\alpha) = (\Psi'(\alpha) - \frac{1}{\alpha})^{-1}.$$

The method of moments estimators for α and β are

$$\tilde{\beta}_n = \frac{\overline{X^2} - (\overline{X})^2}{\overline{X}}$$

and

$$\tilde{\alpha}_n = \frac{(\overline{X})^2}{\overline{X^2} - (\overline{X})^2}.$$

With the help of delta method and the functions

$$g_1(\overline{X}, \overline{X^2}) = \frac{(\overline{X})^2}{\overline{X^2} - (\overline{X})^2}$$

and

$$g_2(\overline{X}, \overline{X^2}) = \frac{\overline{X^2} - (\overline{X})^2}{\overline{X}}$$

we find

$$\sqrt{n}\begin{bmatrix} \tilde{\alpha} - \alpha \\ \tilde{\beta} - \beta \end{bmatrix} \to N(0, \Sigma(\theta))$$

with $\Sigma(\theta) = B\Lambda B^t$ where

$$B = \begin{bmatrix} \frac{2(1+\alpha)}{\beta} & \frac{-1}{\beta^2} \\ \frac{-(2\alpha+1)}{\alpha} & \frac{1}{\alpha\beta} \end{bmatrix}$$

and Λ is the covariance matrix

$$\Lambda = \begin{bmatrix} \alpha\beta^2 & 2\beta^3\alpha(\alpha+1) \\ \frac{2}{\beta}\beta^3\alpha(\alpha+1) & 2\beta^4\alpha(\alpha+1)(2\alpha+3) \end{bmatrix}.$$

After few calculations we obtain

$$\Sigma(\theta) = B\Lambda B^t = \begin{bmatrix} 2a(1+\alpha) & -2\beta(1+\alpha) \\ -2\beta(1+\alpha) & \frac{\beta^2(3+2\alpha)}{\alpha} \end{bmatrix}.$$

From (23) the optimal asymptotic variance of $\sqrt{n}\hat{\alpha}$ when the method of moments estimator for β is available from an independent sample is

$$V_2(\alpha) = (\Psi'(\alpha) - \frac{1}{2\alpha})^{-1}.$$

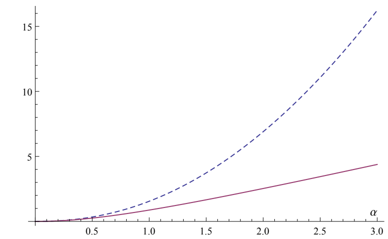

Fig. 1. Graphs of asymptotic variances V_1 (dotted line) and V_2 (solid line).

The Figure 1 shows the relation between the two variances and as it is expected the V_2 is always better compared to V_1.

Next in order to produce the O.E.F. (30), we produced two independent simulated samples $\underline{x}$ and $\underline{y}$ of size $n = 50$. For the finding of the MLE of α from the sample $\underline{x}$ we find the root of the function

$$f(\alpha) = \overline{\ln}(x) - \ln(\overline{x}) + \ln(\alpha) - \Psi(\alpha) \tag{31}$$

In order to write the O.E.F. (30), we find the method of moments estimator for β from the second sample $\underline{y}$ which is $\tilde{\beta} = 3.087$. Thus the (30) becomes

$$J^o_{\text{opt}}(\alpha) = \Psi(\alpha) + \frac{\tilde{\beta}}{2\alpha}\overline{x} - \overline{\ln}(x) + \ln(\tilde{\beta}) - \frac{\tilde{\beta}^2}{2} \tag{32}$$

The next Figure 2 shows the graph of the f and J^o_{opt} estimating functions. From (31) the MLE estimator of a α is $\hat{\alpha} = 1.629$ while the E.F. (32) utilizing

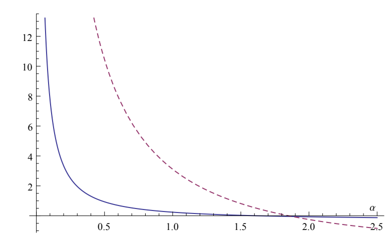

Fig. 2. Graphs the estimating functions J^o_{opt} (dotted line) and f (solid line).

the estimator for β from the second sample gives as an estimation for α the value $\hat{\alpha}^* = 1.851$ which is closer to the true value $\alpha = 2$.

References

1. V.P. Bhapkar. Sufficiency, ancillarity and information in estimating functions. *Estimating Functions.* Edited by Godambe, V. P. Oxford Science Publications, Oxford. 241-254, 1991.
2. P. J. Bickel, C. A. J. Klaassen, Y. Ritov and J. A. Welner. *Efficient and adaptive estimation for semiparametric models.* John Hopkins University Press, Baltimore, 1993.

3. V. P. Godambe.An optimum property of the regular maximum likelihood estimator. *Ann. Math. Statist*, 1208–12, 1960.
4. V. P. Godambe. Onancillarity and Fisher information in the presence of a nuisance parameter. *Biometrika*, *71*, 626–629, 1984.
5. N. L. Johnson, S. Kotz and N. Balakrishnan. *Continuous Univariate Distributions* . Vol. 1, 2nd Ed. John Wiley, New York, 1994.
6. E. I. Lehmann.*Theory of point estimation*. John Wiley and Sons, Inc. New York, 1983.
7. K. Liang. On information and ancillarity in the presence of a nuisance parameter. *Biometrika*, 607–612, 1983.
8. B.Lindsay.Nuisance parameters, mixture models, and the efficiency of partial likelihood estimators. *Phil. Trans. R. Soc. A*. 639–665, 1980.
9. J. Neyman. Optimal asymptotic tests of composite hypothesis. *The Harald Crammer volume*. (Grenander ed.) 213–234, 1959.
10. C. Stein. Efficient non parametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* 187–195. University of California Press, Berkeley, 1956.