CHARACTERIZATIONS INVOLVING CONDITIONAL EXPECTATIONS BASED ON A FUNCTIONAL DERIVATIVE APPROACH

Rosa E. Lillo and Miguel Martín

00-12

WORKING PAPERS

Universidad Carlos III de Madrid

Working Paper 00-12
Statistics and Econometrics Series (6)
February 2000

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 624-9849

CHARACTERIZATIONS INVOLVING CONDITIONAL EXPECTATIONS BASED ON A FUNCTIONAL DERIVATIVE APPROACH.

Rosa E. Lillo and Miguel Martín*

Abstract

We introduce a notion of the derivative with respect to a distribution function, not relating necessarily to probability, which generalizes the concept of the derivative as proposed by Lebesgue (1973). The differential calculus required to solve the linear differential equation involved in this notion of the derivative is included in the paper. The definition given here may also be considered as the inverse operator of a modified notion of the Riemmann--Stieltjes integral. Both this unified approach and the results of differential calculus allow us to characterize distributions in terms of three different types of conditional expectations. In applying these results, a test of goodness-of-fit is also indicated. Finally, two characterizations of a general Poisson process are included, based on conditional expectations. Specifically, a useful result for the homogeneous Poisson process is generalized to a general context.

Keywords: Functional derivative; h-mean lifetime (deathtime) function; general Poisson processes, characterization results.

*Lillo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-91-624.98.57, Fax: 34-91-624.98.49, e-mail: lillo@est-econ.uc3m.es; Martín, Departamento de Estadística e Investigación Operativa, Universidad Complutense de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain, e-mail: martin@mat.ucm.es.

Characterizations involving conditional expectations based on a functional derivative approach

R. E. Lillo *
Dpt. of Statistics and Econometrics
Carlos III University, Madrid

M. Martín
Dpt. of Statistics and OR
Complutense University, Madrid

Abstract

We introduce a notion of the derivative with respect to a distribution function, not relating necessarily to probability, which generalizes the concept of the derivative as proposed by Lebesgue (1973). The differential calculus required to solve the linear differential equation involved in this notion of the derivative is included in the paper. The definition given here may also be considered as the inverse operator of a modified notion of the Riemmann–Stieltjes integral. Both this unified approach and the results of differential calculus allow us to characterize distributions in terms of three different types of conditional expectations. In applying these results, a test of goodness–of–fit is also indicated. Finally, two characterizations of a general Poisson process are included, based on conditional expectations. Specifically, a useful result for the homogeneous Poisson process is generalized to a general context.

1 Introduction

Lebesgue (1973) introduced the notion of the derivative with respect to a function of bounded variation in (a, b) with $-\infty \leq a < b \leq \infty$ as

$$\frac{df(t)}{\alpha(t)} = \lim_{h \to 0} \frac{f(t+h) - f(t)}{\alpha(t+h) - \alpha(t)} \quad \text{for} \quad a < t < b \tag{1}$$

With this definition, Lebesgue (1973) showed that the relation between the Riemann–Stieltjes integral related to α and this notion of the derivative is the same as that existing between the Riemman integral

*Rosa E. Lillo
Dpto. de Estadística y Econometría, Universidad Carlos III
Madrid, 126. 28903 Getafe, Madrid. Spain
E–mail: rlillo@est-econ.uc3m.es

and the usual derivative in the case $\alpha(t) = t$; that is, both operators, derivative and integral, are inverse. Within this framework, this encourages study of differential equations based on the concept of the derivative given in (1). For instance, a first–order differential equation (*α–equation*) might be expressed as follows:

$$\frac{df(t)}{d\alpha(t)} = g\left[t, \alpha(t), f(t)\right] \quad \text{for} \quad a < t < b \tag{2}$$

To develop this idea, let us consider the following problem. Let T be a random variable taking values in (a, b) with a distribution function F and a distribution tail $\bar{F} = 1 - F$. Let us assume furthermore that $E[T]$ is finite. Define, for $a < t < b$,

$$m(t) = E\left[T \mid T > t\right] = \frac{\int_t^b x dF(x)}{\bar{F}(t)} \tag{3}$$

m is called the *mean lifetime function.* This immediately raises the question of how to find the distribution F, starting from knowledge of m itself. In simples cases, the procedure of finding F is familiar from elementary calculus. For example, if the distribution is absolutely continuous, F is obtained by differentiating in (3); that is, by evaluating $dm(t)/dt$. Inspired by normal differential calculus, we propose the idea of replacing dx with dF in order to obtain $dm(t)/dF$. Speaking without complete scientific rigour and extending the classical rules of differentiation, $dm(t)/dF$ would be reduced to:

$$\frac{dm(t)}{dF(t)} = \frac{-t + m(t)}{\bar{F}(t)} \tag{4}$$

Consequently, solving the differential equation involved as if it were conventional; that is, considering $F(t) = t$, the distribution F might be derived as follows:

$$\bar{F}(t) = \exp^{-\int_a^t \frac{dm(x)}{m(x)-x}} \quad \text{for all } a < t < b \tag{5}$$

Obviously, equation (4) is a particular case (linear) of (2). The normal approach to solving (2) consists of interpreting it as an integral equation, using a heuristic argument and then giving adequate meaning to the integral chosen. This approach is usually considered in the context of *stochastic differential equations.* Some simple problems have a form similar form to (2). For example, the known *Growth equation,* (see Karling and Taylor (1975), pp. 354)

$$dX = X\, dB \tag{6}$$

where B is the Brownian Motion. In the above reference, $X(t) = \exp[B(t)]$ is given as a pathwise solution of (6) [with initial condition $X(0) = 1$] which is found in an informal, non-mathematical manner. It is also shown that this solution is the one required when the equation (6) is interpreted as a Stratanovic stochastic differential equation. On the other hand, the solution $Y(t) = \exp[B(t) - t/2]$ agrees with the Ito solution when (6) is interpreted in the sense of Ito. The methodology developed in this approach starts by explicitly formalizing the notion of the *integral.*

In this work, we propose a different approach based on previously formalizing a concept of the derivative which will be useful in solving our problem. However, this is a meaningless approach within the general context of ***stochastic differential equations*** since $df(t)/d\alpha(t)$ usually depends on dt. After introducing the definition of the ***functional derivative***, we will use methods and concepts relating to ordinary differential equations, but applied in this case to the definition of the derivative given here. We should point out that analysis of the existence of solutions and detailed research into procedures for solving (2) will not be included here. We will include only the arguments and developments required to formalize mathematically the problem of characterization of the mean lifetime functions stated in (4). Indeed we will study a more general problem, a problem of characterization in terms of conditional events of the form $(s < T < t)$, which will be denoted as ***bilateral characterization***.

Turning to the example that motivated this work; that is, the characterization of distributions by mean lifetime functions, it should be observed that both Definition 1 and the usual definition of the Riemman–Stieltjes integral are insufficient to achieve the objective mapped out in this paper. The main reason for this is that if F or m are discontinuous functions, the discontinuity points of F and m will be the same. For this reason, in the context of characterizing distributions in terms of conditional expectations, it has been assumed in the literature that the function of distribution is absolutely continuous and, therefore, that classical differential calculus may be involved, applied as the inverse operator of the Lebesgue integral. It is, therefore, necessary to consider modified definitions of derivative and integral if a general approach is sought, without restrictions on F.

Hence our main objective in this paper is to give a perfectly plausible definition of the functional derivative, in the same sense as that provided by Lebesgue, such that the formula 5 may be derived by differentiation. We should point out that once known, the solution to (3), (5) might be obtained by analytical methods, without appealing to a new concept of the derivative as may be seen in Lillo and Martín (1997). However we are interested in formalizing these ideas and procedures, since a more general characterization problem of a conditional nature, and easily solved by these methods will be treated here. Moreover, this differential calculus may be applied in other contexts and to other problems. On the other hand, considering a modified definition of the Riemann–Stieltjes integral, both operators, the functional derivative introduced here and the integral, would undergo process inversion, thus guaranteeing a unified approach in relation to the applications. Section 2 is devoted to introducing this new approach of a derivative compatible with the integral of the equation (3) and (5). An extended subsection is also included, with properties and results relating to previously defined operators.

Advances in the characterization of probability distributions over the last two decades have brought new dimensions to the area of research, and have generated considerable interest among researchers into both Probability and Statistics. Characterization properties are of potential importance in various areas, such as Reliability, Statistical Inference, Stochastic Processes, Bayesian Methods and Information

Theory. Recent results of characterization require conditions to be applied to the distribution function of the random variable. A detailed panorama of the wide interest in the applicability of characterizations in these areas may be found in Galambos (1978), Gupta (1984), Meilijson (1972), Nagaraja and Nevzorov (1997), Rao and Shanbhag (1986), Shanbhag (1970). In Section 3, we characterize *completely h–mean lifetime functions* using the calculus described in Section 2. The definition of *h–mean lifetime function* is a generalization of Definition 3 in the following sense: let h be a strictly increasing function from the interval (a, b) to $\mathbb{R}$, such that the set of continuity points of F is included in the set of continuity points of h; that is, is F–continuous. Define, for $a < t < b$,

$$m_h(t) = E\left[h(T) \mid T > t\right] \quad \text{for } a < t < b \tag{7}$$

then, m_h is called the h–mean lifetime function. Hamdan (1972), Kotlarski (1973) and Kotlarski and Hinds (1975) used this conditional mean to characterize some continuous probabilities. We say *completely* since any assumptions regarding the function distribution will be required in our results. However, in the literature Dealing with this problem, it is often assumed that the distribution function is absolutely continuous.

Some definitions Relating to conditional expectations of random variables are given, and subsequently, theorems of characterization of distributions based on such conditional means are shown. By way of application of a new concept called *bilateral conditional mean*, we propose a test of goodness–of–fit based on the general technique of the *chi–squared test for goodness of fit*. A new approach aimed at deriving the exponential distribution from a process with periodic memory is also characterized in Subsection 3.2. Finally, Section 4 is devoted to proving two theorems of characterization of a general Poisson process. One of them is a generalization of the theorem of Watanabe (1964) for the homogeneous Poisson process. The other is based on the conditional mean of the waiting time until the next arrival.

2 Functional derivative

Obviously, definition of a functional derivative suggests differentiation with respect to a function; that is, the idea of replacing the variable of differentiaton x for a function α. The simplest definition of the derivative with respect a function was introduced by Lebesgue (1973). However, this definition is not plausible, for example, when the function of differentiation is the distribution function related to a discrete random variable. For this reason, we may add some definitions and notation. Before starting the calculus, we should point out that the approach introduced here is focused on diffentiating with respect to distribution functions, since this is sufficient to solve our initial problem. Nevertheless, ideas and results may be extended to the class of bounded variation functions.

Consider an open interval (a, b), $-\infty \leq a < b \leq \infty$ on the real line. Let α be a non–decreasing, right–continuous real function defined in (a,b). From here on, we will assume that α is fixed. Let $\mathcal{F}$ denote the class of real functions defined in (a,b). For $g \in \mathcal{F}$, define

$$\begin{aligned} \mathcal{C}(g) &= \{\, t \,/\, t \in (a,b),\ g \text{ is continuous at } t \,\} \\ \mathcal{D}(g) &= \{\, t \,/\, t \in (a,b),\ g \text{ is discontinuous at } t \,\} \end{aligned}$$

Definition 1 *A point $t \in (a,b)$ is said to be an increasing point of f iff for every $h > 0$ such that for $(t-h, t+h) \subset (a,b)$, we have $f(t-h) < f(t) < f(t+h)$. The set of all such t is denoted by $\mathcal{I}(f)$.*

Definition 2 *Given $f \in \mathcal{F}$, $t \in \mathcal{I}(\alpha)$, the $\alpha^{(1)}$-derivative of f at t is defined as*

$$D_{\alpha}^{(1)} f(t) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{\alpha(t+h) - \alpha(t)} \tag{8}$$

If $t + h \in (a,b)$ and the limit of (8) exists and is finite, f is said to be $\alpha^{(1)}$-differentiable at t.

This definition has been considered by Lebesgue (1973) to establish a unified approach between the Riemann–Stieltjes integral and derivatives. Consequently, in reference to our introductory example described in (3), we have

$$\text{If } t \notin \mathcal{I}(F) \Rightarrow t \notin \mathcal{I}(m_h) \Rightarrow m_h \text{ is not } F^{(1)}\text{–differentiable at } t$$

Hence, Definition 8 is not sufficient to achieve the aim of this work based on differentiating m_h with respect to F.

Definition 3 *A point $t \in (a,b)$ is said to pertain to the support of function f iff for every $h > 0$ such that for $(t-h, t+h) \subset (a,b)$, we have $f(t+h) - f(t-h) > 0$. The set of all such t is denoted by $\mathcal{S}(f)$.*

Another concept of the derivative is now considered in relation to the support of α.

Definition 4 *Given $f \in \mathcal{F}$, $t \in \mathcal{S}(\alpha)$, the $\alpha^{(2)}$-derivative of f at t is defined as*

$$D_{\alpha}^{(2)} f(t) = \lim_{h \to 0} \frac{f(t+h) - f(t-h)}{\alpha(t+h) - \alpha(t-h)} \tag{9}$$

If $(t-h, t+h) \subset (a,b)$ and the limit of (9) exists and is finite, f is said to be $\alpha^{(2)}$-differentiable at t.

It should be remembered that we seek a derivative such that the Riemann–Stieltjes integral may be considered as its inverse process. In this context, Definition 4 is not sufficient to generalize the Barrow's rule although α be continuous. For example, let $(a,b) = (-1,1)$. Let $\alpha(t) = t$ be

$$f(t) = \begin{cases} t+1 & \text{for} \quad t \in (-1,0] \\ 1-t & \text{for} \quad t \in (0,1) \end{cases}$$

Obviously, f is $\alpha^{(2)}$–differentiable for all $t \in (a,b)$, but

$$\int_t^s f(u)d\alpha(u) \neq D_\alpha^{(2)} f(s) - D_\alpha^{(2)} f(t) \quad \text{if} \quad -1 < t < 0 < s < 1$$

Therefore, neither Definition 2 for discontinuity points, nor Definition 4 for continuity points are satisfactory, for which reason we introduce the following,

Definition 5 *Given $f \in \mathcal{F}$, we say that f is $\alpha^{(3)}$-differentiable iff*

- *f is $\alpha^{(1)}$-differentiable for all $t \in \mathcal{C}(\alpha) \cap \mathcal{I}(\alpha)$*
- *f is $\alpha^{(2)}$-differentiable for all $t \in \mathcal{S}(\alpha) \cap \mathcal{D}(\alpha)$*

Then, $D_\alpha^{(3)}$ denotes this derivative.

Remark 1 There are obvious implications with respect to these definitions: f is $\alpha^{(1)}$–differentiable in $(a,b) \Rightarrow f$ is $\alpha^{(3)}$–differentiable in $(a,b) \Rightarrow f$ is $\alpha^{(2)}$–differentiable in (a,b).

Let us suppose that f is $\alpha^{(3)}$–differentiable for all $t \in \mathcal{S}(\alpha)$. $D_\alpha^{(3)}(t)$ may be extended over interval (a,b) in a variety of ways. Note that if $(c,d) \notin \mathcal{S}(\alpha)$, the Riemann–Stieltjes integral,

$$\int_a^d f(u)d\alpha(u) = \int_a^c f(u)d\alpha(u)$$

Bearing this fact in mind, and for the reasons given in the Introduction, we choose the following derivative extension:

Definition 6 *A function f is said to be $\bar{\alpha}^{(3)}$-differentiable for all $t \in (a,b)$ iff*

$$\bar{D}_\alpha^{(3)} f(t) = \begin{cases} D_\alpha^{(3)} f(t) & \text{if } \; t \in \mathcal{S}(\alpha) \\ 0 & \text{if } \; t = a, \; a \notin \mathcal{S}(\alpha) \\ D_\alpha^{(3)} f(t_s) & \text{if } \; t \notin \mathcal{S}(\alpha) \text{ where } t_s = \sup\{u \,/\, u \leq t, \; u \in \mathcal{S}(\alpha)\} \end{cases} \tag{10}$$

This extension is not essential but it will allow us to express some results in a simple way. For simplicity, we will use the notation $D_\alpha = \bar{D}_\alpha^{(3)} f(t)$.

2.1 Rules and properties of differentiation

Let α be a non–decreasing, right–continuous real function defined in (a,b). Now, we introduce two classes of functions

$$\mathcal{D}_\alpha^{(i)} \;\; = \;\; \{f \in \mathcal{F} \,/\, f \text{ is } \alpha^{(i)}\text{–differentiable in } (a,b)\}\,, \; i = 1,2,3$$

$$\begin{aligned}\mathcal{D}^{(i)} &= \{f \in \mathcal{F} \,/\, f \text{ is } \alpha^{(i)}\text{–differentiable in } (a,b) \text{ with } \alpha(t) = t\}, \; i = 1,2,3 \\ \mathcal{D}_\alpha &= \{f \in \mathcal{F} \,/\, f \text{ is } \bar{\alpha}^{(3)}\text{–differentiable in } (a,b)\}\end{aligned}$$

The following basic rules are derived using procedures similar To those applied to the usual derivate. For this reason we will omit the proof. Let $f, g \in \mathcal{D}_\alpha^{(i)}$ be.

1. $D_\alpha^{(i)}(f+g)(t) = D_\alpha^{(i)} f(t) + D_\alpha^{(i)} g(t)$
2. $D_\alpha^{(i)}(fg)(t) = g(t) D_\alpha^{(i)} f(t) + f(t) D_\alpha^{(i)} g(t)$
3. If $g(t) > 0$ for all $t \in (a,b)$
$$D_\alpha^{(i)}\left(\frac{f}{g}\right)(t) = \frac{g(t)D_\alpha^{(i)} f(t) - f(t) D_\alpha^{(i)} g(t)}{g(t)^2}$$
4. $D_\alpha^{(i)}(\log f)(t) = f^{-1} D_\alpha^{(i)} f(t)$
5. If $f \in \mathcal{D}^{(i)}$, $g \in \mathcal{D}_\alpha^{(i)}$, then $f \circ g \in \mathcal{D}_\alpha^{(i)}$ and
$$D_\alpha^{(i)}(f \circ g)(t) = D_\alpha^{(i)} g(t) D^{(i)} f(g(t)) \tag{11}$$

Theorem 1 *If $f \in \mathcal{D}_\alpha^{(2)}$ and $D_\alpha^{(2)} f(t) = 0$ for all $t \in (a,b)$, then f is constant in (a,b).*

Proof: The theorem will follow if, for any interval (c,d) in which the hypotheses hold, $f(c) = f(d)$. Given ϵ, with each point $t \in (c,d)$ is associated an interval $I_t = (t - h_t, t + h_t)$ (arbitrarily small) for which

$$|f(s) - f(s')| < \epsilon|\alpha(s) - \alpha(s')|, \; \forall\, s, s' \in I_t \text{ such that } s < t < s'$$

The set of intervals $\{I_t\}_{t \in [c,d]}$ is an open covering of the compact interval $[c,d]$. Then, we may consider a finite subcovering $\{I_{t_i}\}$, $i = 1, \dots, n$. Assume that $t_1 < t_2 < \cdots < t_n$, and for all i, let $t_{i,i+1}$ verify $t_{i,i+1} \in I_{t_i} \cap I_{t_{i+1}}$ and $t_i < t_{i,i+1} < t_{i+1}$. Considering $t_0 = c$ and $t_{n+1} = d$, we arrive at

$$\begin{aligned}|f(d) - f(c)| &\le \sum_{i=0}^{n} |f(t_i) - f(t_{i,i+1})| + \sum_{i=0}^{n} |f(t_{i,i+1}) - f(t_{i+1})| \\ &\le \epsilon\,(\alpha(d) - \alpha(c))\end{aligned} \tag{12}$$

Since ϵ is arbitrarily small, $f(c) = f(d)$. ■

Remark 2 *Theorem 1 is also valid if α is of bounded variation on (a,b). Then, inequality 12 is rewritten as*

$$|f(d) - f(c)| \le \epsilon T(c,d)$$

where $T(c,d)$ is the total variation of α in $[c,d]$.

2.2 A modified Riemann–Stieltjes integral

Definition 7 *Let J denote a closed linear interval $[t, s]$; let N be any finite set of numbers, say $t_0, t_1, \dots, t_n$, such that*

$$t = t_0 < t_1 < t_2 < \cdots < t_n = s$$

Then N is said to be a net over J, and the closed intervals $[t_{r-1}, t_r]$ $(r = 1, 2, \dots, n)$ are called its cells. The gauge of N, denoted as $g(N)$, is given by

$$g(N) = \max\{t_r - t_{r-1},\ r = 1, 2, \dots, n\}$$

Consider α to be once more a non–decrasing, right–continuous real function, as defined in (a, b).

Definition 8 *Let $f \in \mathcal{F}$ be a real function defined in an interval (a, b). Let N be a net over $[t, s] \subset (a, b)$. Then form the Riemann sum*

$$R_N(f) = \sum_{r=0}^{n-1} [\alpha(t_{r+1}) - \alpha(t_r)]\, f(t_{r+1}) \tag{13}$$

Let $N_1, \dots, N_q, \dots$ be a sequence of nets or partitions such that $g(N_q) \to 0$. If the sequence $R_{N_q}(f)$ has a limit as $q \to \infty$, which is independent of the choice of the sequence N_q [provided only that $g(N_q) \to 0$], then the limit is called the $\hat{\alpha}$–integral of the function f [over interval [t,s]], and we write

$$\widehat{\int_t^s} f(u)d\alpha(u) = \lim_{d(N)\to 0} R_N(f) \tag{14}$$

Remark 3 Note that if we put in (13) $f(t_r)$ instead of $f(t_{r+1})$, the Riemann sums involved lead to the stochastic integral of Ito, (see Karlin and Taylor (1975)).

We keep the symbol $\int_t^s f(u)d\alpha(u)$ to denote the usual Riemann-Stieltjes integral of f with respect α. The existence of the integral can be proved if suitable assumptions are made about functions f and α, the most natural of which are that f is continuous and α monotonic (or of bounded variation). An integral defined in this manner is such that *if f is integrable $d\alpha$ (in the Riemann–Stieltjes sense) then f and α are nowhere simultaneously discontinuous*, (see Kestelman (1959)). However, the example considered in the introduction of this paper reveals that $m_h(t)$ and $F(t)$ are simultaneously discontinuous. Then, m_h is not a Riemann–Stieltjes integrable with respect to $F(t)$. This fact motivated Definition 8 in which function f in the Riemann sums is always evaluated at the lower point of intervals $[t_r, t_{r+1}]$ instead of at any point $\xi \in [t_r, t_{r+1}]$.

On the other hand, it is easy to see that if α is a non–decreasing, right–continuous real function defined in (a, b), there exists a continuous function α_1 and a step–function α_2 such that

$$\alpha = \alpha_1 + \alpha_2 \tag{15}$$

Relation between both definitions of integral is given in the following result.

Lemma 1 *Let there be a function $f \in \mathcal{F}$ such that $\mathcal{C}(f) \subset \mathcal{C}(\alpha)$, then*

$$\widehat{\int_t^s} f(u)d\alpha(u) = \int_t^s f(u)d\alpha_1(u) + \sum_{u \in \mathcal{D}(\alpha)} f(u)\left[\alpha_2(u) - \alpha_2(u-0)\right] \tag{16}$$

Lemma 2 *Let α^* be the measure induced by α in $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ denotes the Borel field. Observe that Lemma 1 indicates that the Lebesgue integral with respect to a measure may be considered as the Riemann–Stieltjes integral given in (14); that is,*

$$\int_t^s f d\alpha^* = \widehat{\int_t^s} f(u)d\alpha(u) \tag{17}$$

We omit the proof, since the issue is straighforward considering (15).

Remark 4 In spite of relation 17, we continue to use $\widehat{\int_t^s}$ to denote the operator integral, since its meaning will be different in this context.

We now include some useful results to show that the derivative introduced before and this modified Riemann–Stieltjes integral are inverse processes.

Theorem 2 *Let α be a non-decreasing, right-continuous real function defined in (a,b) and let $f \in \mathcal{F}$ such that $\mathcal{C}(\alpha) \subset \mathcal{C}(f)$ and $\hat{\alpha}$-integrable in (a,b). Then,*

$$F(t) = \widehat{\int_a^t} f(u)d\alpha(u), \ \forall\, t \in (a,b)$$

Besides, $F \in \mathcal{D}_\alpha$ and $D_\alpha F(t) = f(t)$.

Proof: If $t \in \mathcal{C}(\alpha) \cap \mathcal{I}(\alpha)$, we have

$$\begin{aligned}
\lim_{h\to 0} \frac{F(t+h) - F(t)}{\alpha(t+h) - \alpha(t)} &= \lim_{h\to 0} \frac{\widehat{\int_t^{t+h}} f(u)d\alpha(u)}{\alpha(t+h) - \alpha(t)} \\
&= \lim_{h\to 0} \frac{\widehat{\int_t^{t+h}} \left(f(u) - f(t)\right) d\alpha(u) + \widehat{\int_t^{t+h}} f(t)d\alpha(u)}{\alpha(t+h) - \alpha(t)} \\
&= f(t)
\end{aligned}$$

If $t \in \mathcal{D}(\alpha) \cap \mathcal{S}(\alpha)$, we have

$$\lim_{h\to 0} \frac{F(t+h) - F(t-h)}{\alpha(t+h) - \alpha(t-h)} = \lim_{h\to 0} \frac{\widehat{\int_{t-h}^{t+h}} f(u)d\alpha(u)}{\alpha(t+h) - \alpha(t-h)}$$

$$= \frac{f(t)\,(\alpha(t) - \alpha(t-0))}{\alpha(t) - \alpha(t-0)}$$
$$= f(t)$$

and finally, if $t \notin \mathcal{S}(\alpha)$ the proof is immediate, bearing in mind Definition 6 ■

Theorem 3 *Let $f \in \mathcal{D}_\alpha$ be such that $\mathcal{C}(\alpha) \subset \mathcal{C}(D_\alpha f)$ in (a,b). Then,*

$$\widehat{\int_a^t} D_\alpha \left[f(u)\right] d\alpha(u) = f(t) - f(a+0) \quad \textit{for all } a < t < b \tag{18}$$

Proof: Define

$$H(t) = \widehat{\int_a^t} D_\alpha \left[f(u)\right] d\alpha(u)$$

¿From Theorem 2, we know that $D_\alpha H(t) = D_\alpha f(t)$ which implies that $D_\alpha(H - f)(t) = 0$ for $t \in (a,b)$. Consequently, the result follows from Theorem 1. ■

As an example of this modified Barrow's rule, let

$$\alpha(t) = \begin{cases} 0 & \text{for} \quad t < 0 \\ 1 & \text{for} \quad t \geq 0 \end{cases} \qquad f(t) = \begin{cases} 0 & \text{for} \quad t < 0 \\ c & \text{for} \quad t \geq 0 \end{cases}$$

Then we have,

$$z(t) = D_\alpha f = \begin{cases} 0 & \text{for} \quad t < 0 \\ c & \text{for} \quad t \geq 0 \end{cases}$$

and

$$\widehat{\int_t^s} z(u) d\alpha(u) = f(s) - f(t)$$

Note that Definitions 2, 4, 5 and 6 may also be considered when α is a bounded variation, right-continuous function with its left–limit in (a,b). However this paper is not focused on the in-depth study of operators but applications. Finally, we will include some useful results to be used later.

Theorem 4 *Let $f \in \mathcal{D}_\alpha$ be and $D_\alpha f > 0$ in (a,b). Then, if α is non-decreasing (non–increasing) in (a,b), f is non-decreasing (non–increasing) in (a,b).*

On the other hand, let $f \in \mathcal{D}_\alpha$ be and $D_\alpha f < 0$ in (a,b). Then, if α is non-decreasing (non–increasing) in (a,b), f is non-increasing (non-decreasing) in (a,b).

The proof is omitted since it may be obtained by standard calculations.

Theorem 5 *Let f, α, A, B be functions in (a, b) such that $C(\alpha) \subset C(B)$ and $C(f) \subset C(A)$. Then, we have*

$$D_\alpha f(t) = \frac{A(t,f)}{B(t,\alpha)} \Longleftrightarrow \widehat{\int_a^t} \frac{1}{B(u,\alpha(u))} d\alpha(u) = \widehat{\int_a^t} \frac{1}{A(u,f(u))} df(u) \tag{19}$$

Proof: Assume that the right–hand part is true, then

$$\begin{aligned} \frac{1}{B(t,\alpha)} &= D_\alpha \left(\widehat{\int_a^t} \frac{1}{A(u,f(u))} df(u) \right) \\ &= D_f \left(\widehat{\int_a^t} \frac{1}{A(u,f(u))} df(u) \right) D_\alpha f(t) \\ &= \frac{1}{A(t,f)} D_\alpha f(t) \end{aligned}$$

Considering

$$\frac{1}{A(t,f)} D_\alpha f(t) = \frac{1}{B(t,\alpha)}$$

the other implication is immediate and the proof is complete ∎

Theorem 6 *The solution of the first–order linear differential equation:*

$$D_\alpha f(t) + A(t,\alpha) f(t) = B(t,\alpha) \tag{20}$$

is determined by

$$f(t) = \frac{1}{\mu(t)} \left[\widehat{\int_a^t} \mu(s) B(s,\alpha(s)) d\alpha(s) + c \right] \tag{21}$$

where

$$\mu(t) = \exp \left\{ \widehat{\int_a^t} A(s,\alpha(s)) d\alpha(s) \right\}$$

Proof: We solve the equation by means of the integrant factor method; that is, we look for $\mu(t)$ verifying:

$$\mu(t) \left[D_\alpha f(t) + A(t,\alpha) f(t) \right] = D_\alpha \left[\mu f \right](t)$$

After the calculations, the following is obtained

$$\frac{D_\alpha \mu(t)}{\mu(t)} = A(t,\alpha) \Rightarrow \mu(t) = \exp \left\{ \widehat{\int_a^t} A(s,\alpha(s)) d\alpha(s) \right\}$$

Then,

$$D_\alpha \left[\mu f \right](t) = \mu(t) B(t,\alpha)$$

¿From which the result easily follows. ∎

3 The h–mean lifetime function

Let T be a random variable with $a < T < b$ where $-\infty \leq a < b \geq \infty$ and a finite mean. Let F be its distribution function. Let h be a strictly increasing, F–continuous function from the interval (a, b) to $\mathbb{R}$. We propose the following definition, where a more general concept of mean lifetime is introduced.

Definition 9 *Consider a random variable T and a function h described above. Then,*

$$m_h(t) = E\,[\,h(T)\,|\,T > t\,]\,, \quad \textit{for} \quad a < t < b \tag{22}$$

is called the h–mean lifetime function (HMLF).

Note that if $h(t) = t$ and if (a, b) is taken to be $[0, \infty)$, we have the mean lifetime function typically used in the *Survival Theory* to describe the aging phenomenon. Specifically, Meilijson (1972) obtains the distribution functions related to positive random variables, with finite means and an infinite essential supremum starting from the mean lifetime function. ¿From here on, $m(t)$ will denote the mean lifetime function related to function $h(t) = t$. A natural question is: what are the conditions necessary and sufficient to guarantee that a function m_h is a HMLF?. This paper gives a fairly complete answer to this question. For any distribution function F, denote $\bar{F} = 1 - F$.

Certain characterizations of probability distributions have been obtained by using properties of conditional expectations. Shanbhag (1970) gave a characterization of exponential distribution in terms of conditional expectations. A generalization of this kind of result is given by Hamdan (1972). In this reference, an HMLF is considered assuming that h is a strictly increasing differentiable function. Kotlarski and Hinds (1975) used the concept of conditional expectation to characterize continuous probability distributions on the real line. However, the distribution functions are required to be absolutely continuous and the functions h differentiable or continuously differentiable, (see Patil et al. (1975) for a quick review). The following result is the widest generalization of this kind of characterizations. The difference in the approach here is that instead of investigating absolutely continuous distributions functions, any distribution on the interval (a, b), $-\infty \leq a < b \leq \infty$ may be characterized in terms of conditional expectations.

Theorem 7 *Let m_h be a function on (a, b) and let h be a strictly increasing, right-continuous function. Then, m_h is an HMLF for some distribution function F if, and only if, the following conditions are fulfilled*

1. *$m_h(t)$ is a right-continuous fuction and $m_h(t) > h(t)$ for all $a < t < b$.*

2. *$m_h(t)$ is a non–decreasing function for all $a < t < b$.*

3. *For all $a < t < b$,*

$$\widehat{\int_a^t} \frac{dm_h(x)}{m_h(x) - h(x)} < \infty$$

4. *Taking the limits to be $t \to b$*

$$\widehat{\int_a^b} \frac{dm_h(x)}{m_h(x) - h(x)} = \infty$$

Moreover, let m_h be a function fulfilling these four properties, then there exists a unique function $\bar{F}$, such that $m_{h\bar{F}} = m_h$, where $m_{h\bar{F}}$ denotes the HMLF related to the distribution function $F = 1 - \bar{F}$. In this case, $\bar{F}$ is determined as

$$\bar{F}(t) = \exp\left\{ -\widehat{\int_a^t} \frac{dm_h(x)}{m_h(x) - h(x)} \right\} \tag{23}$$

Proof: Firstly, we assume that m_h is the HMLF related to some distribution function F. ¿From the definition of HMLF given in (22) and considering that h is strictly increasing, it is immediate that $m_h(t) > h(t)$, and from Theorem 2 in section 2.1, it may be deduced that an $HMLF$ is F–continuous and therefore, a right–continuous function. Since m_h can be stated in connection with $\bar{F}$, i.e,

$$m_h(t) = \frac{\widehat{\int_t^b} h(x) dF(x)}{\bar{F}(t)} \tag{24}$$

condition 2 may be obtained using property 4 of the derivatives.

$$D_F m_h(t) = \frac{-h(t)\bar{F}(t) + \widehat{\int_t^b} h(x) dF(x)}{\bar{F}(t)^2} \tag{25}$$

Since h and F are increasing, (25) is greater than zero and then, $m_h(t)$ is non-decreasing. To verify the third condition, observe that

$$D_{m_h} \bar{F}(t) = -\frac{\bar{F}(t)}{m_h(t) - h(t)} \tag{26}$$

Then, we can write

$$\widehat{\int_a^t} \frac{D_{m_h} \bar{F}(x)}{\bar{F}(x)} dm_h(x) = \widehat{\int_a^t} - \frac{dm_h(x)}{m_h(x) - h(x)}$$

Applying Theorem 3 to the left–hand term, we obtain equation (23)

$$\bar{F}(t) = \exp\left\{ -\widehat{\int_a^t} \frac{dm_h(x)}{m_h(x) - h(x)} \right\}$$

Consequently, the third and fourth conditions are immediate. Now, we will prove the sufficient conditions. To this end, we will show that function $\bar{F}$ defined in (23) determines a distribution function in (a, b) with finite mean. Obviously $\bar{F}(a + 0) = 1$ and $\bar{F}(b - 0) = 0$. Using once more derivative Theorem 2, it is easy to see that $\bar{F}$ is right–continuous. On the other hand, we may prove that $\bar{F}(t)$ is a non–increasing function of t. To this end, it is sufficient to differentiate with respect to m_h.

$$D_{m_h}\bar{F}(t) = -\exp\left\{-\widehat{\int_a^t} \frac{dm_h(x)}{m_h(x)-h(x)}\right\}\frac{1}{m_h(t)-h(t)} \tag{27}$$

Since m_h is a non–decreasing function and (27) is negative, $\bar{F}$ is non–increasing. Finally, the expected value of $h(T)$ associated with the distribution derived from $\bar{F}$ is $m_h(a)$ which is finite and then, $E[T]$ is also finite due to the properties of function h. Now we have only to show that $m_{h\bar{F}} = m_h$. By equation (26), we know that $m_{h\bar{F}}$ verifies

$$D_F m_{h\bar{F}}(t) = \frac{m_{h\bar{F}}(t)-h(t)}{\bar{F}(t)} \tag{28}$$

Observe that (28) is a linear differential equation whose unique solution under the condition $m_{h\bar{F}}(a) = E[h(T)] = m_h(a)$ is

$$m_{h\bar{F}}(t) = \frac{\widehat{\int_t^b} h(x)dF(x)}{\bar{F}(t)}$$

which is derived using Theorem 6. ¿From (27), we have that m_h verifies (28). Since the solution is unique, we have that $m_{h\bar{F}} = m_h$ and the proof is complete. ■

Let T be a positive random variable, with finite mean and infinite essential supremum.

$$\text{Define, for}\quad 0 \le t < \infty,\quad \mu(x) = E\left[T-t\,|\,T>t\right]$$

μ is called the *mean residual lifetime function*, MRLF.

Corollary 1 *Let μ be a positive function on $[0,\infty)$. Then, μ is an MRLF If, and only if,*

1. $\mu(t)$ *is a right-continuous fuction for all* $t \ge 0$.
2. $m(t) = \mu(t) + t$ *is a non-decreasing function for all* $t > 0$.
3. $\int_0^t \frac{dx}{\mu(x)} < \infty$ *for all* $t \ge 0$.
4. $\int_0^\infty \frac{dx}{\mu(x)} = \infty$

Moreover, let μ be a function fulfilling these four properties, then there exists a unique function $\bar{F}$ such that $\mu_{\bar{F}} = \mu$, where $\mu_{\bar{F}}$ denotes the MRLF related to the distribution function $F = 1-\bar{F}$. In this case, $\bar{F}$ is determined as

$$\bar{F} = 1 - F(t) = \frac{\mu(0)}{\mu(t)}\exp\left\{-\int_0^t \frac{dx}{\mu(x)}\right\}$$

Proof: Observe that

$$m(t) = t + \mu(t)$$

is the MLF associated with T. This fact leads to the conclusion of proof applying Theorem 7. ■

Remark 5 Corollary 1 has been proven by Lillo and Martín (1997) using a direct development, without considering the differential calculus introduced here and applied in a Bayesian context.

In the same way, we also define the concept of *h–mean deathtime function* as a conditional mean by evens $(T \leq t)$

Definition 10 *Consider a random variable T taking values on the interval (a, b) and finite mean. Let F be its distribution function and consider a strictly increasing, F–continuous function h from the interval (a, b) onto $\mathbb{R}$. Then,*

$$d_h(t) = E\left[h(T) \,|\, T \leq t\right], \quad \text{for } a < t < b \tag{29}$$

is called the h–mean deathtime function (HMDF).

Applying an approach similar to that used to prove Theorem 7, we can also characterize the h–mean deathtime functions, as written in the following theorem:

Theorem 8 *Let d_h be a function from (a, b). Then, d_h is an HMDF for some distribution function F if, and only if, the following conditions are satisfied*

1. *$d_h(t)$ is a right–continuous fuction and $d_h(t) < h(t)$ for all $a < t < b$.*
2. *$d_h(t)$ is a non–decreasing function for all $a < t < b$.*
3. *For all $a < t < b$,*
$$\widehat{\int_t^b} \frac{dd_h(x)}{h(x) - d_h(x)} < \infty$$
4. *Taking the limit to be $t \to a$*
$$\widehat{\int_a^b} \frac{dd_h(x)}{h(x) - d_h(x)} = \infty$$

Moreover, let d_h be a function fulfilling these four properties, then there exists a unique distribution function F, such that $d_{hF} = d_h$, where d_{hF} denotes the HMDF related to the distribution function F which is determined as

$$F(t) = \exp\left\{-\int_t^b \frac{dd_{hF}(x)}{h(x) - d_{hF}(x)}\right\} \tag{30}$$

We omit the proof since it is based on the same arguments as the proof for Theorem 7. As a consequence of Theorem 7, results of characterization of probability distributions on the real line, in terms of conditional expectations, may be obtained. The first is a generalization of a kind of characterization given by Hamdan (1972).

Theorem 9 *Let h be a strictly increasing, right-continuous function from the interval (a,b) to $[0,\infty)$ and let c be a positive constant satisfying $a < c < b$. The random variable T has its cumulative distribution given by*

$$F(t) = \begin{cases} 0 & t \leq a \\ 1 - e^{-\frac{h(t)}{h(c)}} & a < t < b \\ 1 & t \geq b \end{cases} \tag{31}$$

if, and only if,

$$\forall\, t \in (a,b),\ m_h(t) = E[h(T) \,|\, T > t] = h(t) + h(c) \tag{32}$$

Proof: ¿From Theorem 7 it is sufficent to verify that m_h is an HMLF. Then, using (23), the distribution function is expressed as

$$\bar{F}(t) = e^{-\widehat{\int_a^t} \frac{dh(x)}{h(c)}} = e^{-\frac{h(t)}{h(c)}}, \quad for\ a \leq t \leq b$$

and the proof is complete. ■

Corollary 2 *If (a,b) is taken to be $(0,\infty)$ and $h(t) = t^d, t \in [0,\infty),\ d > 0,$ Theorem 9 gives a characterization of the Weibull distribution*

$$\bar{F}(t) = e^{-\frac{t^d}{c^d}}, \quad for\ t \leq 0$$

If in addition we have $d = 1$, the Weibull distribution reduces to the exponential distribution. In this case, the condition necessary and sufficient for T to be an exponential distribution with mean a is that

$$m(t) = t + a \quad \forall\, t > 0 \tag{33}$$

Continuing with the idea of characterizing distributions in terms of conditional expectations, we introduce another definition of conditional expectation based on a bidimensional function.

Definition 11 *Consider a random variable T taking values on the interval (a,b), finite mean and distribution function F. Consider a function strictly increasing, F-continuous function h from the interval (a,b). Then,*

$$\mu_h(t,s) = E\left[\,h(T)\,|\, t < T < s\,\right], \quad for\ a < t < s < b \tag{34}$$

is called the h-bilateral conditional mean function (HBCMF).

We will see that the HBCMF is a useful tool to characterize probability functions involving a fit–test with application in Statistical Inference. Firstly, three theorems of characterization of these functions are given without proof, since they are easily followed using Theorem 7 and Theorem 8.

Theorem 10 *Let $\mu_h(t,s)$ be a function from $(a,b) \times (a,b)$. Then, $\mu_h(t,s)$ is an HBCMF for some distribution function F if, and only if,*

1. *For all $t \in (a,b)$, the limit*
$$\lim_{s \uparrow b} \mu_h(t,s) = \mu_{h,b}(t)$$
exists and $\mu_{h,b}$ is an HMLF related to the interval (a,b).

2. *If F_b is the distribution function associated to $\mu_{h,b}$, It verifies,*
$$\mu_h(t,s) = \frac{\widehat{\int_t^s} h(x) dF_b(x)}{F_b(s) - F_b(t)} \tag{35}$$
whenever $F_b(t) < F_b(s)$.

Theorem 11 *Let $\mu_h(t,s)$ be a function from $(a,b) \times (a,b)$. Then, d_h is an HBCMF for some distribution function F if, and only if,*

1. *For all $s \in (a,b)$, the limit*
$$\lim_{t \downarrow a} \mu_h(t,s) = \mu_{h,a}(s)$$
exists and $\mu_{h,a}$ is an HMDF related to the interval (a,b).

2. *If F_a is the distribution function associated to μ_a, It verifies equation (35).*

Theorem 12 *Let $\mu_h(t,s)$ be a function from $(a,b) \times (a,b)$. Then, d_h is an HBCMF for some distribution function F if, and only if, $\mu_b(t)$ is an HMLF, $\mu_a(s)$ is a HMDF and moreover, $F_b(x) = F_a(x)\ \forall\, x \in (a,b)$.*

Proof: It sufficient to consider that if m_F and d_F are the HMLF and the HMDF related the same distribution F, the following relation is verified

$$d_F(x) = \frac{\mu_h}{F(x)} - \frac{\bar{F}(x)}{F(x)} m_F(x)$$

where μ_h denotes the expected value of $h(T)$. ∎

As an example of the usefulness of this theorem, bilateral conditional means that may be written as linear combination of t and s will be characterized.

Theorem 13 *Let $h(x) = x$ be. A function $\mu(t,s) = \lambda t + (1-\lambda)s$, $a < t < s < b$, $0 < \lambda < 1$ is a bilateral conditional mean if, and only if, $\lambda = 1/2$. In this case, the associated distribution is uniform over the interval (a,b).*

Proof: Taking the limits to be $s \uparrow b$ and $t \downarrow a$, we have

$$\mu_b(t) = \lambda t + (1-\lambda)b \tag{36}$$

$$\mu_a(s) = \lambda a + (1-\lambda)s \tag{37}$$

According to Theorem 7 and Theorem 8, μ_b is an MLF and μ_a is an MDF, respectively with the following associated distribution functions:

$$\bar{F}_b(t) = \left(\frac{b-t}{b-a}\right)^{\frac{\lambda}{1-\lambda}}, \; F_a(s) = \left(\frac{s-a}{b-a}\right)^{\frac{1-\lambda}{\lambda}} \tag{38}$$

¿From Theorem 12, we require $\bar{F}_b(t) = 1 - F_a(t)$ for $t \in (a,b)$. By straightforward calculations, it is easy to prove that this is possible if, and only if, $\lambda = 1/2$. In this case, the distribution involved is the uniform distribution over the interval (a,b). ∎

3.1 Goodness–of–fit test

We propose a test of goodness-of-fit based on the notion of the *bilateral conditional mean* and inspired by the general technique of the *chi-square test for goodness of fit*. The method of the test proposed here proceeds as follows. Suppose we are sampling from a d.f. $F(x)$ with support for the interval (a,b), (which may depend on parameter(s) θ). Divide the range of the distribution into k mutually exclusive and exhaustive intervals, say $I_1, I_2, \ldots, I_k$, where $I_i = [t_{i-1}, t_i)$ and $t_0 = a$, $t_k = b$. Each interval has a conditional bilateral mean and variance of containing an r.v with d.f. $F(x)$,

$$\mu_i = \mu(t_{i-1}, t_i) = \frac{\widehat{\int_{t_{i-1}}^{t_i}} x dF(x)}{F(t_i) - F(t_{i-1})}, \; i = 1, 2, \ldots, k$$

$$\sigma_i^2 = \sigma^2(t_{i-1}, t_i) = \frac{\widehat{\int_{t_{i-1}}^{t_i}} (x-\mu_i)^2 dF(x)}{F(t_i) - F(t_{i-1})}, \; i = 1, 2, \ldots, k$$

Each sample value falls into exactly one of the intervals. Let $N_1, N_2, \ldots, N_k$ be the respective observed numbers of observations of the sample in the intervals $I_1, I_2, \ldots, I_k$ and let $N = N_1 + \cdots + N_k$ be the sample size. Remember that the *chi-square test for goodness of fit* is based on the vector $\mathbf{N} = (N_1, N_2, \ldots, N_k)$ has a multinomial distribution and mainly evaluates a measure of the diference between the theoretical frequencies and observed frequencies that the data might well have come from $F(x)$. However once the vector **N** is evaluated, the information provided by the sample values is not borne in mind. The test introduced here attempts to incorporate this information in the following way.

Firstly, we have to introduce more notation. The random sample will be written more conveniently as X_{ij}, where i, $i = 1, \ldots, k$ denotes the interval in which the observation falls and j, $j = 1, \ldots, N_i$ denotes the order in the sample *i–th*. It is obvious that

$$\begin{aligned} E_F\left[X_{ij} \mid X_{i,j} \in I_i\right] &= \mu_i,\ i = 1, \ldots, k,\ j = 1, \ldots, N_i \\ Var_F\left[X_{ij} \mid X_{i,j} \in I_i\right] &= \sigma_i^2,\ i = 1, \ldots, k,\ j = 1, \ldots, N_i \end{aligned}$$

Let $\bar{X}_i$ be the sample mean related to the interval I_i. Then,

$$\frac{\sqrt{N_i}(\bar{X}_i - \mu_i)}{\sigma_i} \xrightarrow{d} N(0,1) \quad \text{as} \quad N_i \to \infty \tag{39}$$

On the other hand, fixed vector **N** and the random variables $\bar{X}_i$, $i = 1, \ldots, k$ are independent and thus,

$$T = \sum_{i=1}^{k} \frac{\sqrt{N_i}(\bar{X}_i - \mu_i)}{\sqrt{k}\sigma_i} \tag{40}$$

is asymtotically distributed $N(0,1)$. Now, we wish to test the hypothesis $H_0 : F(x) = F_0(x)$ where any parameters in $F_0(x)$ are completely specified. First compute μ_i and σ_i^2, $i = 1, \ldots, k$. If H_0 is true, then intuitively we expect $\bar{X}_i \simeq \mu_i$. Thus T further away from zero indicates data less compatible with the claimed null distribution. Hence for large n a level α test is given by the critical region;

$$\mathcal{R} = \left\{T < -z_{\frac{\alpha}{2}}\right\} \cup \left\{T > z_{\frac{\alpha}{2}}\right\}$$

Remark 6 Suppose now that the number of intervals may change with sample size. We denote it by k_n. Note that an interesting problem in this context would be analizying conditions over k_n to ensure that the asymtotic power of the test is one.

3.2 Exponential with periodic memory

Consider that the life time of a system is governed by a random phenomenon with period d. In terms of conditional expectations, the associated MLF should verify

$$m(t) = t + m\left(t - d\left[\frac{t}{d}\right]\right) \tag{41}$$

where $[x]$ denotes the integer–part of x; that is, we want a mean lifetime function determined by its values in an interval $(0, d)$. In the next result, we will characterize the distributions associated with this form of MLF.

Theorem 14 *An MRL verifies (41) if, and only if, there exists a number* p, $0 < p < 1$ *and a distribution function* $\bar{F}_d(t)$ *defined over* $(0, d)$*, such that the distribution function related to (41) is*

$$\bar{F}(t) = p^{\left[\frac{t}{d}\right]} \left\{ (1-p)\bar{F}_d \left(t - d \left[\frac{t}{d} \right] \right) + p \right\} \tag{42}$$

Proof: Let F be the distribution function related to (41). From Theorem 7, it is easy to see that $\bar{F}(t)$ satisfies:

$$\bar{F}(t) = \left(\bar{F}(d)\right)^{\left[\frac{t}{d}\right]} \bar{F} \left(t - d \left[\frac{t}{d} \right] \right)$$

Taking,

$$\begin{aligned} p &= \bar{F}(d) \\ \bar{F}_d(x) &= \frac{\bar{F}(x) - p}{1-p} \quad \forall\, x \in (0, d) \end{aligned}$$

the neccesary condition is complete, since $\bar{F}_d$ determines a distribution over $(0, d)$. To prove the condition sufficient, define $\bar{F}(t)$ as in (42). Then, $1 - \bar{F}(t)$ determines a distribution function over $(0, \infty)$. Using again Theorem 7 and by straightforwad calculations, the mean lifetime function associated to this distribution verifies condition (41). ■

Remark 7 This result is interesting since if we take the limit to be $d \to 0$, the limit distribution is exponential with mean $m(0)$, which introduces a new approach to explain the exponential distribution.

4 Characterization of General Poisson Processes

Consider an arrival process N; that is, for any $\omega \in \Omega$, the mapping $t \to N_t(\omega)$ is non–decreasing, increases by jumps only, and is right continuous with $N_0(\omega) = 0$. In this section, we focus on characterizing general Poisson processes. First, recall (see Cinlar (1975)) that a general Poisson process has associated an expectation function a meaning

$$a(t) = E\left[N_t\right],\ t > 0 \tag{43}$$

By the definition of a general Poisson process, a is a non–decreasing right continuous function. Let τ_t be the waiting time until the next arrival after t. Following the idea of Corollary 2 applied to the exponential distributuion in (33), we want first to characterize a general Poisson Process in terms of the expected value of τ_t. A second characterization is involved based on the expected number of arrivals during an interval of time.

Theorem 15 *Let a be a non-decreasing, right continuous function and let N be an arrival process. Then, N is a general Poisson process if, and only if,*

$$E[\tau_t \mid N_u,\, u \leq t] = \widehat{\int_0^\infty} e^{-[a(t+s)-a(t)]} ds = A(t) \tag{44}$$

Proof: It is easy to see that N is a non–stationary Poisson process with an expectation function a if, and only if, the distribution of τ_t, with the past history $\{N_u;\, u \leq t\}$ of the process known to t is such as

$$\bar{F}_t(x) = P[\tau_t > x \mid N_u\, u \leq t] = e^{-[a(t+x)-a(t)]} \tag{45}$$

Thus, the neccesary condition is obviously involved. To prove the condition sufficient, we define

$$B_t(s) = E_t[\tau_t \mid \tau_t > s]$$

where $E_t[\cdot]$ denotes *expectation when the past history $\{N_u;\, u \leq t\}$ is known up to t.* By straighforward calculations, we can write

$$\begin{aligned} B_t(s) &= \frac{E_t[\tau_t I_{\tau_t > s}]}{P[\tau_t > s \mid N_u,\, u \leq t]} \\ &= \frac{E_t[E_{t+s}[\tau_t I_{\tau_t > s}]]}{P[\tau_t > s \mid N_u,\, u \leq t]} \\ &= \frac{E_t[(s + A(t+s))\, I_{\tau_t > s}]}{P[\tau_t > s \mid N_u,\, u \leq t]} \\ &= s + A(t+s) \end{aligned}$$

Since $B_t(s)$ is a mean lifetime function, Theorem 7 allows us to obtain the distribution of τ_t as

$$\bar{F}_t(x) = \frac{A(t)}{A(t+x)} e^{-\int_0^x \frac{ds}{B_t(s)-s}} \tag{46}$$

Taking,

$$c(t) = \widehat{\int_t^\infty} e^{-a(u)} du$$

we have

$$\widehat{\int_0^x} \frac{ds}{B_t(s) - s} = \ln\left(\frac{c(t)}{c(t+x)}\right)$$

Since $A(t) = e^{a(t)}c(t)$, (46) can be rewritten as (45) and the proof is complete. ■

¿From a practical point of view, checking to see if the axioms given in the definition of a general (non–stationary) Poisson process hold for a particular process may be quite difficult. The following theorem considerably reduces the checks involved.

Theorem 16 *With the same assumptions as in Theorem 15, $N = \{N_t; t \geq 0\}$ is a general Poisson process with an expectation function a if, and only if,*

$$E\left[N_{t+s} - N_t \mid N_u\, u \leq t\right] = a(t+s) - a(t)$$

Proof: Let F_t be the distribution function related to the random variable τ_t. With this notation, we can write:

$$\begin{aligned} E_t\left[N_{t+s} - N_t\right] &= \widehat{\int_0^s} E_t\left[N_{t+s} - N_t \mid \tau_t = x\right] dF_t(x) \\ &= \widehat{\int_0^s} \left(1 + E_{t+x}\left[N_{t+s} - N_{t+x} \mid \tau_t = x\right]\right) dF_t(x) \end{aligned} \tag{47}$$

In view of the condition of the statement of the theorem, we have the following equation:

$$a(t+s)\bar{F}_t(s) - a(t) - F_t(s) = -\widehat{\int_0^s} a(t+x)dF_t(x) \tag{48}$$

Taking the derivative with respect to F_t in (48),

$$\bar{F}_t(s)D_{F_t}a(t+s) = 1$$

¿From Theorem 5, we have

$$a(t+s) - a(t) = -\ln(1 - F_t(s))$$

Then, the distribution of τ_t verifies equation (45) and N is a general Poisson process. The proof of the necessary condition is omitted since it is immediate. ■

Remark 8 It may be noted that if $a(t) = t$, we have also characterized the stationary Poisson process. Specifically, Theorem 16 is the generalization of Watanabe 's theorem (1964) for stationary processes.

5 Final Remarks

In view of the unified approach taken in this paper between the modified Riemann–Stieltjes integral and the functional derivative, it will interesting to study other typical problems of the modern theories of differentiation, such as differential equations with order greater than one. Some work is currently in progress in this direction, inspired mainly by its application to the characterization of distributions defined in subsets of $\mathbb{R}^n$ with $n \geq 1$. We believe that the concept of ***bilateral conditional mean*** in the univariate case might be generalized to a multivariate setup and provide a a test of goodness–of–fit based on the development of Subsection 3.1.

References

Adler, R.J. and Brown, L.D. (1986) Tail behaviour for suprema of empirical processes. *Ann. Probab.*, **14**, 1–30.

Arnold, B.C., (1997) Characterizations involving conditional specification. *J. Statis. Plan. Inf.*, **63**, 117–131.

Balkema, A.A. and Haan, L., (1974) Residual life time at great age. *Ann. Prob.*, **2**, 792–804.

Berger, J.O., (1980) *Statistical Decision Theory*, New York–Heidelberg–Berlin, Springer.

Burkill, J.C., (1958) *The Lebesgue Integral*, Cambrigde University Press.

Cinlar, E. (1975) *Introduction to Stochastic Processes*. Pretince–Hall, Inc.

Chung, K. L., (1974) *A course in Probability Theory*, Academic Press, INC.

Galambos, J. and Kotz, S. (1978) *Characterizations of probability distributions*. Lecture Notes in Mathematics No. 675, Springer, Berlin.

Goldstein, M., (1975) Uniqueness relations for linear posterior expectation. *J. Roy. Statist. Soc. Ser B*, **37**, 402–405.

Gupta, R. C. (1984), Relations between order statistics and record values and some characterizations results. Uniform mixtures via posterior means. *J. Appl. Probab.*, **21**, 425–430.

Gupta, R.C. and Sankaran, P.G. (1998), Bivariate equilibrium distribution and its applications to Reliability. *Commun. Statist.-Theory Meth*, **27**(2), 385–394.

Gupta, A. K. and Wesolowski, J. (1997), Uniform mixtures via posterior means. *Ann. Inst. Statist. Math.*, **1**, 171–180.

Hamdan, M.A. (1972) On a characterization by conditional expectations. *Technometrics*, **14**, 497–499.

Karlin, S. and Taylor, H.M. (1975) *A second course in Stochastic Processes*. Academic Press, INC London.

Kestelman, H., (1959) *Modern theories of integration*, Dover publications, INC. New York.

Kotlarski, I. (1973). On a characterization of some distributions by conditional expectations. *Sanhkhya A*, **34**, 461–467.

Kotlarski, I. and Hinds, W.E. (1975). Characterization theorems using conditional expectations. *J. App. Prob.*, **12**, 400–406.

Lebesgue, H. (1973) *L'intégration*. Chelsea publishing, New York.

Lillo, R.E. and Martín, M. (1997) Characterization of Bayes estimators of the uniform scale parameters. *Statistics & Decisions*, **17**, 31–47.

Meilijson, I., (1972) Limiting properties of the mean residual lifetime function. *Ann. Math. Statist.*, **43**, 354–357.

Patil, G. P, Kotz, S. and Ord, J. K., (1975) *Statistical Distribution in Scientific Work, Vol. 3*, 221–230. Reidel Publishing Company, Dordrecht–Holland.

Rao, C.R. and Shanbhag, D.N. (1986) Recent results on characterization of probability distributions: a unified approach through extensions of Deny's theorem. *Adv. Appl. Probab.*, **18**, 660–678.

Sreehari, M. and Kalamkar, V.A. (1997) Modelling some stationary Markov processes and related characterizations. *J. Statist. Plan. Inf.*, **63**, 363–375.

Shanbhag, D. N. (1970) The characterizations for exponential and geometric distributions. *J. Amer. Stat. Assoc.*, **65**, 1256–1259.

Shilov, G. E. and Gurevich B. L., (1977) *Integral, measure and derivative: a unified approach*, Dover publications, INC. New York.

Watanabe, S. (1964) On discontinuous additive functionals and Lévy measures of a Markov process. *Japan J. Math.* **34**, 53–70.