Combinatorics, Probability and Computing (2001) 10, 417–434. © 2001 Cambridge University Press
DOI 10.1017/S0963548301004825 Printed in the United Kingdom

Characterization Results Based on a Functional Derivative Approach

R. E. LILLO[1] and M. MARTIN[2]

[1] Dpto. de Estadística y Econometria, Universidad Carlos III de Madrid, Madrid, 126. 28903 Getafe, Madrid, Spain
(e-mail: lillo@est-econ.uc3m.es)

[2] Dpto. de Estadística e I.O.
Universidad Complutense, Madrid, Spain
(e-mail: mmartin@mat.ucm.es)

Received 19 April 1999; revised 7 February 2001

We introduce a notion of the derivative with respect to a function, not necessarily related to a probability distribution, which generalizes the concept of derivative as proposed by Lebesgue [14]. The differential calculus required to solve linear differential equations using 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 Riemann–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. Specifically, a useful result for the homogeneous Poisson process is generalized.

1. Introduction

Lebesgue [14] introduced the notion of the derivative with respect to a function of bounded variation in (a, b), with $-\infty \leqslant a < b \leqslant \infty$, by

$$\frac{df(t)}{d\alpha(t)} = \lim_{h \to 0} \frac{f(t+h) - f(t)}{\alpha(t+h) - \alpha(t)}, \qquad \text{for } a < t < b. \tag{1.1}$$

With this definition, Lebesgue [14] showed that the relationship between the Riemann–Stieltjes integral related to $\alpha(t)$ and this notion of the derivative is the same as that existing between the Riemann integral and the usual derivative in the case $\alpha(t) = t$: that is, both operators, derivative and integral, are inverse. This framework encourages the study of differential equations based on the concept of derivative given in (1.1). For instance, a

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

first-order differential α-equation might be stated as follows:

$$\frac{df(t)}{d\alpha(t)} = g\,[t, \alpha(t), f(t)]\,, \qquad \text{for } a < t < b. \tag{1.2}$$

The normal approach to solving (1.2) 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 to (1.2), for example, the well-known *growth equation* (see Karling and Taylor [8], p. 354)

$$dX(t) = X(t)\,dB(t), \tag{1.3}$$

where $B(t)$ is Brownian motion. In the above reference, $X(t) = \exp[B(t)]$ is given as a pathwise solution of (1.3) (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 equation (1.3) 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 (1.3) is interpreted in the sense of Ito. The methodology developed in this work starts by explicitly formalizing the notion of the *integral*.

We propose a different approach based on formalizing a concept of the functional derivative such that, 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 these two operators. However, this is a meaningless approach within the general context of *stochastic differential equations* since $df(t)/d\alpha(t)$ usually depends on dt. In order to solve (1.2), we will use methods related to ordinary differential equations, but applied to the definition of the derivative given here. We should point out that analysis of the existence of solutions and detailed research into procedures will not be included in this work. Section 2 is devoted to introducing these two operators.

A field where stochastic differential equations appear is the determination of probability distributions by conditional expectations. Advances in characterization properties are of potential importance in various areas, such as reliability, statistical inference, stochastic processes, Bayesian methods, and information theory. Recent results 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 references [1, 4, 5, 16, 17, 19, 20]. Without making any assumption on the distribution function, Kotz and Shanbhag [12] determine the distribution of an r.v. T in terms of conditional expectations of the type $E\left[h(T) \mid T \geqslant t\right]$. In Section 3, we complete the results given in [12] by determining *completely* the distribution of T in terms of $m_h(t) = E\left[h(T) \mid T > t\right]$, under the same assumptions on h as in [12], but using the calculus described in Section 2, and by characterizing mathematically the set of functions $m_h(t)$. We use other definitions related to conditional expectations of random variables, and subsequently, we prove theorems of characterization based on such conditional means.

As an application of the characterization results, we propose a test of goodness of fit based on the general technique of the *chi-squared test for goodness of fit*. We also

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

determine a new approach aimed at deriving the exponential distribution from a process with periodic memory. As another application of the stochastic differential equation (1.2), Section 4 is devoted to proving two theorems of characterization of a general Poisson process. One is a generalization of the theorem of Watanabe [21] for the homogeneous Poisson process and the other is based on conditional means.

2. Functional derivatives

The simplest definition of the derivative with respect to a function was introduced by Lebesgue [14]. However, this definition is not plausible, for example, when the function of differentiation is the distribution function related to a discrete random variable. Since our problem is focused on differentiating with respect to distribution functions, we may add some definitions and notation. Before starting the calculus, we should point out that ideas and results of this section may be extended to the class of bounded variation functions.

Consider an open interval (a,b), $-\infty \leqslant a < b \leqslant \infty$. Let $\alpha(t)$ be a non-decreasing, right-continuous real function defined in (a,b). From here on, we will assume that $\alpha(t)$ is fixed. Let $\mathscr{F}$ denote the class of real functions defined in (a,b). For $f \in \mathscr{F}$, let $\mathscr{C}(f)$, $(\mathscr{D}(f))$ be the set of continuity (discontinuity) points of f in (a,b). Let $\mathscr{I}(f)$ be the set of increasing points of f, that is, $t \in \mathscr{I}(f)$ if, for every $h > 0$ such that $(t-h, t+h) \subset (a,b)$, $f(t-h) < f(t) < f(t+h)$. Finally, let $\mathscr{S}(f)$ be the support of f. Then, $t \in \mathscr{S}(f)$ if, for every $h > 0$ such that $(t-h, t+h) \subset (a,b)$, $f(t+h) - f(t-h) > 0$.

Definition 1. Let $f \in \mathscr{F}$, $t \in \mathscr{I}(\alpha)$, and $t + h < b$. If the limit

$$\lim_{h \to 0} \frac{f(t+h) - f(t)}{\alpha(t+h) - \alpha(t)} \tag{2.1}$$

exists and is finite, $f(t)$ is said to be $\alpha^{(1)}$-*differentiable at* t. We denote it by $D_{\alpha}^{(1)} f(t)$.

This definition has been considered by Lebesgue [14] to establish a unified approach between the Riemann–Stieltjes integral and derivative. However, it is not sufficient to achieve the aim of this work since it is not possible to differentiate with respect to a discrete distribution function. Hence, another concept of the derivative is now considered in relation to the support.

Definition 2. Let $f \in \mathscr{F}$, $t \in \mathscr{S}(\alpha)$, and $(t-h, t+h) \subset (a,b)$. If the limit

$$\lim_{h \to 0} \frac{f(t+h) - f(t-h)}{\alpha(t+h) - \alpha(t-h)} \tag{2.2}$$

exists and is finite, $f(t)$ is said to be $\alpha^{(2)}$-*differentiable at* t. We denote it by $D_{\alpha}^{(2)} f(t)$.

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 2 is not sufficient to generalize Barrow's rule although $\alpha(t)$ is continuous. For example, let $(a,b) =$

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

$(-1,1)$. Let $\alpha(t)=t$, and let

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

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

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

Therefore, neither Definition 1 for discontinuity points nor Definition 2 for continuity points are satisfactory; for that reason we introduce the following.

Definition 3. Given $f\in\mathscr{F}$, we say that f is $\alpha^{(3)}$-*differentiable* if and only if $f(t)$ is $\alpha^{(1)}$-differentiable for all $t\in\mathscr{C}(\alpha)\cap\mathscr{I}(\alpha)$, and $f(t)$ is $\alpha^{(2)}$-differentiable for all $t\in\mathscr{S}(\alpha)\cap\mathscr{D}(\alpha)$. We then write $D_\alpha^{(3)}f(t)$ to denote this derivative.

There are obvious implications with respect to these definitions: if $f(t)$ is $\alpha^{(1)}$-differentiable in (a,b), then $f(t)$ is $\alpha^{(3)}$-differentiable in (a,b), which implies that $f(t)$ is $\alpha^{(2)}$-differentiable in (a,b).

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

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

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

Definition 4. A function $f(t)$ is said to be $\bar{\alpha}^{(3)}$-*differentiable for all* $t\in(a,b)$ if and only if

$$\bar{D}_\alpha^{(3)}f(t)=\begin{cases} D_\alpha^{(3)}f(t), & \text{if } t\in\mathscr{S}(\alpha),\\ 0, & \text{if } t=a,\ a\notin\mathscr{S}(\alpha),\\ D_\alpha^{(3)}f(t_s), & \text{if } t\notin\mathscr{S}(\alpha) \text{ where } t_s=\sup\{u/u\leqslant t,\ u\in\mathscr{S}(\alpha)\}.\end{cases}\tag{2.3}$$

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 $D_\alpha f(t)=\bar{D}_\alpha^{(3)}f(t)$. Now, we define a modified Riemann–Stieltjes integral.

Definition 5. Let J denote a closed linear interval $[t,s]$; let N be any finite set of numbers, say $t=t_0<t_1<t_2<\cdots<t_n=s$. Then N is said to be a *net* over J, and the closed intervals $[t_{r-1},t_r]$ $(r=1,2,\ldots,n)$ are called its *cells*. The *gauge* of N, denoted as $g(N)$, is given by

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

Definition 6. Let $f\in\mathscr{F}$ be a real function defined in an interval (a,b), and $\alpha(t)$ a non-decreasing, right-continuous real function. Let N be a net over $[t,s]\subset(a,b)$. Then

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

form the *Riemann sum*

$$R_N(f) = \sum_{r=0}^{n-1} [\alpha(t_{r+1}) - \alpha(t_r)]\, f(t_{r+1}). \tag{2.4}$$

Let $N_1, \ldots, N_q, \ldots$ 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 $f(t)$* (over interval $[t, s]$), and we write

$$\widehat{\int_t^s} f(u)d\alpha(u) = \lim_{g(N)\to 0} R_N(f). \tag{2.5}$$

Remark 1. Note that, if we put $f(t_r)$ in (2.4) instead of $f(t_{r+1})$, the Riemann sums involved lead to the stochastic integral of Ito (see Karlin and Taylor [8]).

We keep the symbol $\int_t^s f(u)d\alpha(u)$ to denote the usual Riemann–Stieltjes integral of $f(t)$ with respect to $\alpha(t)$. The existence of this integral can be proved if suitable assumptions are made about functions $f(t)$ and $\alpha(t)$, the most natural of which are that $f(t)$ is continuous and $\alpha(t)$ monotonic (or of bounded variation). An integral defined in the Riemann–Stieltjes sense implies that $f(t)$ and $\alpha(t)$ are nowhere simultaneously discontinuous (see Kestelman [9]). However, we will consider examples where both f and α are simultaneously discontinuous. Then, $f(t)$ is not Riemann–Stieltjes-integrable with respect to $\alpha(t)$. This fact motivated Definition 6, in which function $f(t)$ in the Riemann sums is always evaluated at the lower point of intervals $[t_r, t_{r+1}]$ instead of at any point in ξ in that interval. On the other hand, it is easy to see that, if $\alpha(t)$ is a non-decreasing, right-continuous real function defined in (a, b), there exists a continuous function $\alpha_1(t)$ and a step-function $\alpha_2(t)$ such that

$$\alpha(t) = \alpha_1(t) + \alpha_2(t), \tag{2.6}$$

which implies the following result.

Lemma 2.1. *Let a function $f \in \mathscr{F}$ be such that $\mathscr{C}(f) \subset \mathscr{C}(\alpha)$. Then*

$$\widehat{\int_t^s} f(u)d\alpha(u) = \int_t^s f(u)d\alpha_1(u) + \sum_{x\in\mathscr{D}(\alpha)} f(x)\,[\alpha_2(x) - \alpha_2(x-0)]\,. \tag{2.7}$$

□

Let α^* be the measure induced by $\alpha(t)$ in $(\mathbb{R}, \mathscr{B})$, where $\mathscr{B}$ denotes the Borel field. Observe that Lemma 2.1 and (2.6) indicate that the Lebesgue integral with respect to a measure may be considered as the Riemann–Stieltjes integral given in (2.5); that is,

$$\int_t^s f\,d\alpha^* = \widehat{\int_t^s} f(u)d\alpha(u). \tag{2.8}$$

In spite of relation (2.8), we continue to use $\widehat{\int_t^s}$ to denote the operator integral, since its meaning will be different in this context.

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Let $\mathscr{D}_\alpha^{(i)}$, $(\mathscr{D}_\alpha)$ be the set of $\alpha^{(i)}$ $(\bar{\alpha}^{(3)})$-differentiable functions in (a,b), $i=1,2,3$. We write $\mathscr{D}^{(i)}$ if $\alpha(t)=t$. The following basic rules are derived using procedures similar to those applied to the usual derivative. For this reason we will omit the proof. It is obvious that all results related to $\mathscr{D}_\alpha^{(3)}$ hold for $\mathscr{D}_\alpha$.

Theorem 2.2. *Let $f,g\in\mathscr{D}_\alpha^{(i)}$.*

(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 $f\in\mathscr{D}^{(i)}$, $g\in\mathscr{D}_\alpha^{(i)}$, then $f\circ g\in\mathscr{D}_\alpha^{(i)}$ and*

$$D_\alpha^{(i)}(f\circ g)(t)=D_\alpha^{(i)}g(t)D^{(i)}f(g(t)). \tag{2.9}$$

□

We now include some auxiliary results that are useful tools in the calculations of the next sections, as well as some theorems to show that the derivative introduced above and this modified Riemann–Stieltjes integral are inverse processes. Their proofs are in the Appendix.

Theorem 2.3. *If $f\in\mathscr{D}_\alpha^{(2)}$ and $D_\alpha^{(2)}f(t)=0$ for all $t\in(a,b)$, then $f(t)$ is constant in (a,b).*

Theorem 2.4. *Let $\alpha(t)$ be a non-decreasing, right-continuous real function defined in (a,b), and let $f\in\mathscr{F}$ be such that $\mathscr{C}(\alpha)\subset\mathscr{C}(f)$ and $\hat{\alpha}$-integrable in (a,b). Consider the function*

$$F(t)=\widehat{\int_a^t} f(u)d\alpha(u),\qquad \text{for all } a<t<b.$$

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

Theorem 2.5 (Barrow's rule). *Let $f\in\mathscr{D}_\alpha$ be such that $\mathscr{C}(\alpha)\subset\mathscr{C}(D_\alpha f)$ in (a,b). Then*

$$\widehat{\int_a^t} D_\alpha(f(u))\,d\alpha(u)=f(t)-f(a+0),\qquad \text{for all } a<t<b. \tag{2.10}$$

Theorem 2.6. *Let $f\in\mathscr{D}_\alpha$ and $D_\alpha f(t)>0$ in (a,b). Then, if $\alpha(t)$ is non-decreasing (non-increasing) in (a,b), $f(t)$ is non-decreasing (non-increasing) in (a,b). On the other hand, let $f\in\mathscr{D}_\alpha$ and $D_\alpha f<0$ in (a,b). Then, if $\alpha(t)$ is non-decreasing (non-increasing) in (a,b), $f(t)$ is non-increasing (non-decreasing) in (a,b).* □

Theorem 2.7. *Let f,α,A,B be functions in (a,b) such that $\mathscr{C}(\alpha)\subset\mathscr{C}(B)$ and $\mathscr{C}(f)\subset\mathscr{C}(A)$. Then we have*

$$D_\alpha f(t)=\frac{A(t,f)}{B(t,\alpha)}\iff\widehat{\int_a^t}\frac{1}{B(u,\alpha(u))}d\alpha(u)=\widehat{\int_a^t}\frac{1}{A(u,f(u))}df(u). \tag{2.11}$$

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Now, we solve a kind of differential equation proposed in (1.2). Its proof is also in the Appendix.

Theorem 2.8. *The solution of the first-order linear differential equation*

$$D_\alpha f(t) + A(t,\alpha)f(t) = B(t,\alpha) \tag{2.12}$$

is determined by

$$f(t) = \frac{1}{\mu(t)}\left[\widehat{\int_a^t}\mu(s)B(s,\alpha(s))d\alpha(s) + c\right], \tag{2.13}$$

where

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

3. The *h*-mean lifetime function

Let T be a random variable with finite mean, and $a < T < b$ where $-\infty \leqslant a < b \leqslant \infty$. Let $F(t)$ be its distribution function. Let $h(t)$ be a real-valued Borel-measurable function from (a,b).

Definition 7. The functions

$$m_h(t) = E\left[h(T) \mid T > t\right], \qquad d_h(t) = E\left[h(T) \mid T \leqslant t\right], \qquad \text{for } a < t < b, \tag{3.1}$$

are called, respectively, the *h-mean lifetime function* (HMLF) and the *h-mean deathtime function* (HMDF).

From here on, $m(t)$ will denote the HMLF related to $h(t) = t$. Note that, if $(a,b) = (0,\infty)$, we have the mean lifetime function typically used in *survival theory* to describe the aging phenomenon. Specifically, Meilijson [16] obtains the distribution functions related to positive random variables, starting from $m(t)$. A natural question is: What are the conditions necessary and sufficient to guarantee that a function $m_h(t)$ is an HMLF? This paper gives a fairly complete answer to this question. For any distribution function $F(t)$, let $\bar{F}(t)$ be $1 - F(t)$.

Certain characterizations of probability distributions have been obtained by using properties of conditional expectations (see references in the Introduction). In [20, 7], an HMLF is considered, assuming that $h(t)$ is a strictly increasing differentiable function. Kotlarski and Hinds [11] used HMLF to characterize continuous probability distributions on the real line, but the distribution functions are required to be absolutely continuous and the functions $h(t)$-differentiable or continuously differentiable (see Patil, Kotz and Ord [18] for a quick review). A representation of the distribution of T in terms of the HMLF, under weaker conditions on h, is given by Kotz and Shanbhag [12]. The following result completes this representation by characterizing the class of HMLFs.

Theorem 3.1. *Let $m_h(t)$ be a function on (a,b) and let $h(t)$ be a real-valued Borel-measurable function such that it is non-decreasing on (a,b) and $h(t) < h(b-)$, for all $t \in (a,b)$. Then,*

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

$m_h(t)$ is an HMLF for some distribution function $F(t)$ if and only if the following conditions are fulfilled.

(1) *$m_h(t)$ is a right-continuous function 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 $m_h(t)$ be a function fulfilling these four properties. Then there exists a unique function $\bar{F}(t)$, such that $m_{h\bar{F}}(t) = m_h(t)$, where $m_{h\bar{F}}(t)$ denotes the HMLF related to the distribution function $F(t) = 1 - \bar{F}(t)$. In this case, $\bar{F}(t)$ is determined as

$$\bar{F}(t) = \exp\left\{-\widehat{\int_a^t} \frac{dm_h(x)}{m_h(x) - h(x)}\right\}. \tag{3.2}$$

Proof. Firstly, we assume that $m_h(t)$ is the HMLF related to some distribution function $F(t)$. From the definition of HMLF given in (3.1) and considering the assumptions made on $h(t)$, it is immediate that $m_h(t) > h(t)$, and from Theorem 2.4 it may be deduced that an HMLF is F-continuous, and therefore a right-continuous function. Since $m_h(t)$ can be stated in connection with $\bar{F}(t)$, that is,

$$m_h(t) = \frac{\widehat{\int_t^b} h(x)dF(x)}{\bar{F}(t)}, \tag{3.3}$$

condition (2) may be obtained by differentiating m_h as follows:

$$D_F m_h(t) = \frac{-h(t)\bar{F}(t) + \widehat{\int_t^b} h(x)dF(x)}{\bar{F}(t)^2}. \tag{3.4}$$

Since (3.4) is greater than zero, $m_h(t)$ is non-decreasing. To verify the third condition, observe that

$$D_{m_h}\bar{F}(t) = -\frac{\bar{F}(t)}{m_h(t) - h(t)}. \tag{3.5}$$

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 2.5 to the left-hand side term, we obtain equation (3.2),

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

and consequently the third and fourth conditions are immediate. Now we will prove the sufficient conditions. To this end, we will show that the function $\bar{F}(t)$ defined in (3.2)

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

determines a distribution function in (a, b) with finite mean. Obviously $\bar{F}(a+0) = 1$ and $\bar{F}(b-0) = 0$. Using Theorem 2.4, it is easy to see that $\bar{F}(t)$ is right-continuous. On the other hand, we may prove that $\bar{F}(t)$ is a non-increasing function of t by differentiation with respect to $m_h(t)$, that is,

$$D_{m_h}\bar{F}(t) = -\exp\left\{-\widehat{\int_a^t} \frac{dm_h(x)}{m_h(x) - h(x)}\right\} \frac{1}{m_h(t) - h(t)}. \tag{3.6}$$

Since $m_h(t)$ is a non-decreasing function and (3.6) is negative, $\bar{F}(t)$ is non-increasing. Finally, the expected value of $h(T)$ associated with the distribution derived from $\bar{F}(t)$ is $m_h(a)$, which is finite, and then $E[T]$ is also finite due to the properties of the function $h(t)$. Now we have only to show that $m_{h\bar{F}}(t) = m_h(t)$. By equation (3.5), we know that $m_{h\bar{F}}(t)$ verifies

$$D_F m_{h\bar{F}}(t) = \frac{m_{h\bar{F}}(t) - h(t)}{\bar{F}(t)}. \tag{3.7}$$

Observe that (3.7) 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 2.8. From (3.6), we have that $m_h(t)$ verifies (3.7). Since the solution is unique, we have that $m_{h\bar{F}}(t) = m_h(t)$ and the proof is complete. □

The determination of the distribution function of an r.v. by other functions, such as the vitality function or the cumulative hazard function, is usually linked with the notion of Lebesgue–Stieltjes measure, in order to apply the Radon–Nikodym theorem (see, for example, [12, 13]). However, the notion of derivative introduced here provides the distribution function of the r.v. and characterizes $m_h(t)$ using arguments typical of differential calculus. In ([12], Proposition 3), it is shown how the d.f. F can be obtained from $E[h(X)|, X \geqslant x]$ whenever h verifies the same assumptions as in Theorem 3.1. With certain obvious alterations to that result, (3.2) can be expressed as

$$\bar{F}(t) = \frac{m_h(a) - h(a)}{m_h(t) - h(t)} \left\{\prod_{t_r \in D_{ax}} g^*(t_r)\right\} \exp\left\{-\int at \frac{dh_c(x)}{m_h(x) - h(x)}\right\}, \tag{3.8}$$

where D_{ax} is the set of discontinuity points of h in the interval (a, x), h_c denotes the continuous part of h in (a, b), and

$$g^*(t_r) = \left\{\frac{m_h(t_r-) - h(t_r)}{m_h(t_r) - h(t_r)}\right\} \frac{m_h(t_r) - h(t_r+)}{m_h(t_r-) - h(t_r-)}.$$

Remark 2. Let T be a positive random variable, with finite mean and infinite essential supremum. Let $\mu(t)$ denote the *mean residual lifetime function* (MRLF), associated with T. Since $\mu(t) = m(t) - t$, a characterization of $\mu(t)$ can be derived as an immediate consequence of Theorem 3.1. Hall and Wellner [6] and Lillo and Martín [15] obtained the

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

same conclusion using a direct development, without considering the differential calculus introduced here, and applied in a Bayesian context.

We can also characterize the h-mean deathtime functions. We omit the proof since it is based on the same arguments as the proof of Theorem 3.1.

Theorem 3.2. *Let $d_h(t)$ be a function from (a,b) and h as in Theorem 3.1. Then, $d_h(t)$ is an HMDF for some distribution function $F(t)$ if and only if the following conditions are satisfied.*

(1) *$d_h(t)$ is a right-continuous function 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 $d_h(t)$ be a function fulfilling these four properties. Then there exists a unique distribution function $F(t)$, such that $d_{hF}(t) = d_h(t)$, where $d_{hF}(t)$ denotes the HMDF associated with the distribution function $F(t)$, which is given by

$$F(t) = \exp\left\{ -\widehat{\int_t^b} \frac{dd_{hF}(x)}{h(x) - d_{hF}(x)} \right\}. \tag{3.9}$$

□

As a consequence of Theorem 3.1, we can extend a kind of characterization of probability distribution given by Hamdan [7].

Theorem 3.3. *Let $h(t)$ 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*

$$F(t) = \begin{cases} 0, & t \leqslant a, \\ 1 - e^{-\frac{h(t)}{h(c)}}, & a < t < b, \\ 1, & t \geqslant b, \end{cases} \tag{3.10}$$

if and only if

$$m_h(t) = h(t) + h(c), \qquad \text{for all } t \in (a,b). \tag{3.11}$$

Proof. From Theorem 3.1 it is sufficient to verify that $m_h(t)$ is an HMLF. Then, using (3.2), the distribution function is expressed as

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

and the proof is complete. □

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Corollary 3.4. *If* $(a,b) = (0,\infty)$ *and* $h(t) = t^d$, $t \geqslant 0$, $d > 0$, *Theorem 3.3 gives a characterization of the Weibull distribution with d.f.*

$$F(t) = 1 - e^{-\frac{t^d}{c^d}}, \qquad \text{for } t \leqslant 0.$$

If in addition we have $d = 1$, *the Weibull distribution reduces to the exponential distribution. In this case,* T *is an r.v. exponentially distributed with mean a if and only if*

$$m(t) = t + a, \qquad \text{for } t > 0. \tag{3.12}$$

□

4. Bilateral conditional mean function

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

Definition 8. Let T be a random variable taking values on the interval (a,b), $E[T] < \infty$, and distribution function $F(t)$. Let $h(t)$ be a real-valued Borel-measurable function in (a,b). Then the function

$$\mu_h(t,s) = E\left[h(T) \mid t \leqslant T \leqslant s\right], \qquad \text{for all } a < t \leqslant s < b, \tag{4.1}$$

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

We will see that the HBCMF is a useful tool in characterizing probability functions involving a test of goodness of fit with applications in statistical inference. Firstly, characterizations of these functions are given. The first result easily follows using Theorems 3.1 and 3.2.

Theorem 4.1. *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* $F(t)$ *if and only if the following conditions are satisfied.*

(1) *For all* $t \in (a,b)$ $(s \in (a,b))$, *the limit*

$$\lim_{s\uparrow b} \mu_h(t,s) = \mu_{h,b}(t), \quad \left(\lim_{t\downarrow a} \mu_h(t,s) = \mu_{h,a}(s)\right),$$

exists and $\mu_{h,b}$ $(\mu_{h,a})$ *is an HMLF (HMDF) associated with the interval* (a,b).

(2) *Let* F_b (F_a) *be the distribution function associated with* $\mu_{h,b}$ $(\mu_{h,a})$. *Taking* $G = F_b$ $(G = F_a)$, *it verifies the equation*

$$\mu_h(t,s) = \frac{\widehat{\int_t^s} h(x)dG(x)}{G(s) - G(t)}, \tag{4.2}$$

whenever $G(t) < G(s)$. □

Theorem 4.2. *The function* $\mu_h(t,s)$ *is an HBCMF for some distribution function* $F(t)$ *if and only if* $\mu_b(t)$ *is an HMLF,* $\mu_a(s)$ *is an HMDF and, moreover,* $F_b(x) = F_a(x)$ *for all* $x \in (a,b)$.

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Proof. It is sufficient to consider that, if m_F and d_F are the HMLF and HMDF associated with the same distribution F, then

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

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

As an example of the usefulness of this theorem, we characterize 'bilateral conditional' means that are a linear combination of t and s.

Theorem 4.3. *Let $h(x) = x$. A function $\mu(t,s) = \lambda t + (1-\lambda)s$, $a < t \leqslant s < b$, $0 < \lambda < 1$ 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

$$\mu_b(t) = \lambda t + (1-\lambda)b, \qquad \mu_a(s) = \lambda a + (1-\lambda)s.$$

According to Theorems 3.1 and 3.2, μ_b is an MLF and μ_a is an MDF, respectively, with the following associated distribution functions:

$$\bar{F}_b(t) = \left(\frac{b-t}{b-a}\right)^{\frac{\lambda}{1-\lambda}}, \qquad F_a(s) = \left(\frac{s-a}{b-a}\right)^{\frac{1-\lambda}{\lambda}}. \tag{4.3}$$

From Theorem 4.2, we require $\bar{F}_b(t) = 1 - F_a(t)$ for $t \in (a,b)$. By straightforward calculation, 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). □

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-squared test for goodness of fit*. The method of the test proposed here proceeds as follows. Suppose we are sampling from a d.f. $F(t)$ 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 $t_0 = a$, $t_k = b$. Consider the conditional bilateral mean and variance

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

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

Each sample value falls into exactly one of the intervals. Let N_i, $i = 1, \ldots, k$, be the number of observations in the intervals I_i and let $N = N_1 + \cdots + N_k$ be the sample size. Remember that the *chi-squared test for goodness of fit* is based on the vector $\mathbf{N} = (N_1, N_2, \ldots, N_k)$ having a multinomial distribution, and it gives a measure of the difference between the theoretical frequencies and observed frequencies. However, once

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

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 order in the ith subsample. It is obvious that

$$E_F\left[X_{ij}|\, X_{i,j} \in I_i\right] = \mu_i, \qquad \mathrm{Var}_F\left[X_{ij}|\, X_{i,j} \in I_i\right] = \sigma_i^2, \quad i = 1,\ldots,k, \quad j = 1,\ldots,N_i.$$

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

$$\frac{\sqrt{N_i}(\bar{X}_i - \mu_i)}{\sigma_i} \xrightarrow{d} N(0,1), \qquad \text{as} \quad N_i \to \infty. \tag{4.4}$$

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

$$T = \sum_{i=1}^{k} \frac{\sqrt{N_i}(\bar{X}_i - \mu_i)}{\sqrt{k}\sigma_i} \tag{4.5}$$

is asymptotically $N(0,1)$-distributed. Now we wish to test the hypothesis $H_0 : F(t) = F_0(t)$. 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 that data are less compatible with the claimed null distribution. Hence, for large n, a level α test is given by the critical region:

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

Remark 3. Suppose now that the number of intervals k_n can change with sample size. Note that an interesting problem in this context would be analysing conditions on k_n to ensure that the asymptotic power of the test is one.

5. Exponential with periodic memory

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

$$m(t) = t + m\left(t - d\left[\frac{t}{d}\right]\right), \tag{5.1}$$

where $[x]$ denotes the integer part of x, that is, the distribution function is determined by the restriction of $m(t)$ to $(0,d)$. The next result characterizes this kind of MLF.

Theorem 5.1. *The MLF $m(t)$ satisfies* (5.1) *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* (5.1) *is*

$$\bar{F}(t) = p^{[\frac{t}{d}]}\left\{(1-p)\bar{F}_d\left(t - d\left[\frac{t}{d}\right]\right) + p\right\}. \tag{5.2}$$

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Proof. Let $F(t)$ be the distribution function associated with (5.1). From Theorem 3.1, 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 $p = \bar{F}(d)$ and

$$\bar{F}_d(t) = \frac{\bar{F}(t) - p}{1 - p}, \qquad \text{for all } 0 < t < d,$$

the necessary condition is complete, since $\bar{F}_d(t)$ determines a distribution over $(0, d)$. Now, define $\bar{F}(t)$ as in (5.2). Then, $1 - \bar{F}(t)$ determines a distribution function over $(0, \infty)$. Again using Theorem 3.1, we have proved that the MLF associated to this distribution verifies (5.1). □

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 explaining the exponential distribution.

6. 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, and it is right-continuous with $N_0(\omega) = 0$. In this section, we focus on characterizing general Poisson processes. First, recall (see Cinlar [3]) that a general Poisson process has an associated expectation function $a(t)$ defined by

$$a(t) = E\left[N_t\right], \qquad \text{for } t > 0. \tag{6.1}$$

By the definition of a general Poisson process, $a(t)$ 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 3.4 applied to the exponential distribution in (3.12), 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 6.1. *Let $a(t)$ 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*

$$E\left[\tau_t \,|\, N_u,\, u \leqslant t\right] = \widehat{\int_0^\infty} e^{-[a(t+s)-a(t)]}\, ds = A(t). \tag{6.2}$$

Proof. It is easy to see that N is a non-stationary Poisson process with an expectation function $a(t)$ if and only if the distribution of τ_t, with the past history $\{N_u;\, u \leqslant t\}$ of the process known to t, satisfies

$$\bar{F}_t(x) = P\left[\tau_t > x \,|\, N_u,\, u \leqslant t\right] = e^{-[a(t+x)-a(t)]}. \tag{6.3}$$

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

$$B_t(s) = E_t\left[\tau_t \,|\tau_t > s\right],$$

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

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

$$B_t(s) = \frac{E_t\left[\tau_t I_{\tau_t > s}\right]}{P\left[\tau_t > s \mid N_u,\ u \leqslant t\right]} = \frac{E_t\left[E_{t+s}\left[\tau_t I_{\tau_t > s}\right]\right]}{P\left[\tau_t > s \mid N_u,\ u \leqslant t\right]}$$

$$= \frac{E_t\left[(s + A(t+s))\, I_{\tau_t > s}\right]}{P\left[\tau_t > s \mid N_u,\ u \leqslant t\right]} = s + A(t+s).$$

Since $B_t(s)$ is an MLF, Theorem 3.1 allows us to obtain the distribution of τ_t as

$$\bar{F}_t(x) = \frac{A(t)}{A(t+x)} e^{-\int_0^x \frac{ds}{B_t(s)-s}}. \tag{6.4}$$

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 $A(t) = e^{a(t)} c(t)$, (6.4) can be rewritten as (6.3), and the proof is complete. □

From a practical point of view, checking to see whether 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 checking involved.

Theorem 6.2. *With the same assumptions as in Theorem 6.1,* $N = \{N_t; t \geqslant 0\}$ *is a general Poisson process with an expectation function* $a(t)$ *if and only if*

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

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

$$E_t\left[N_{t+s} - N_t\right] = \widehat{\int_0^s} E_t\left[N_{t+s} - N_t \mid \tau_t = x\right] dF_t(x)$$

$$= \widehat{\int_0^s} \left(1 + E_{t+x}\left[N_{t+s} - N_{t+x} \mid \tau_t = x\right]\right) dF_t(x). \tag{6.5}$$

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

$$a(t+s)\bar{F}_t(s) - a(t) - F_t(s) = -\widehat{\int_0^s} a(t+x) dF_t(x). \tag{6.6}$$

Taking the derivative with respect to F_t in (6.6),

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

From Theorem 2.7, we have

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

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Then, the distribution of τ_t verifies equation (6.3) and N is a general Poisson process. The proof of the necessary condition is omitted since it is immediate. □

Remark 4. It may be noted that, if $a(t) = t$, we have also characterized the stationary Poisson process. Specifically, Theorem 6.2 is the generalization of Watanabe's theorem [21] for stationary processes.

7. Final remarks

In view of the unified approach taken in this paper between the modified Riemann–Stieltjes integral and the functional derivative, it will be 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 on subsets of $\mathbb{R}^n$, $n \geqslant 1$. We believe that the concept of *bilateral conditional mean* in the univariate case might be generalized to a multivariate setup, and provide a test of goodness of fit based on the development of Section 4.

Appendix

Proof of Theorem 2.3. The theorem will follow if, for any interval (c,d) in which the hypotheses hold, $f(c) = f(d)$. Given ϵ, for each point $t \in (c,d)$ there is an interval $I_t = (t - h_t, t + h_t)$ (arbitrarily small) for which

$$|f(s) - f(s')| < \epsilon |\alpha(s) - \alpha(s')|, \qquad \text{for all } s, s' \in I_t \text{ such that } s < t < s'.$$

The set of intervals $\{I_t,\ c \leqslant t \leqslant d\}$, is an open covering of the compact interval $[c,d]$. Then, we may consider a finite subcovering $\{I_{t_i}\}$, $i = 1, \ldots, n$. 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

$$|f(d) - f(c)| \leqslant \sum_{i=0}^{n} |f(t_i) - f(t_{i,i+1})| + \sum_{i=0}^{n} |f(t_{i,i+1}) - f(t_{i+1})| \leqslant \epsilon\,(\alpha(d) - \alpha(c))\,.$$

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

Proof of Theorem 2.4. If $t \in \mathscr{C}(\alpha) \cap \mathscr{I}(\alpha)$, we have

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

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

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

and finally, if $t \notin \mathscr{S}(\alpha)$, the proof is immediate, bearing in mind Definition 4. □

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

Proof of Theorem 2.5. Define

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

From Theorem 2.4, 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 2.3. □

Proof of Theorem 2.7. Assume that the right-hand side 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. □

Proof of Theorem 2.8. We solve the equation by means of the integrating factor method: that is, we look for $\mu(t)$ verifying

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

After calculation, 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,

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

from which the result easily follows. □

Acknowledgements

The authors would like to thank the referees for their valuable comments.

References

[1] Arnold, B. C. (1997) Characterizations involving conditional specification. *J. Statist. Plan. Inf.* **63** 117–131.
[2] Chung, K. L. (1974) *A Course in Probability Theory*, Academic Press.
[3] Cinlar, E. (1975) *Introduction to Stochastic Processes*, Prentice-Hall.
[4] Galambos, J. and Kotz, S. (1978) *Characterizations of Probability Distributions*, Vol. 675 of *Lecture Notes in Mathematics*, Springer, Berlin.

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press

[5] 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.
[6] Hall, W. J. and Wellner, J. A. (1981) Mean residual life. In *Statistics and Related Topics* (Csörgö, *et al.* eds), North-Holland, Amsterdam, pp. 169–184.
[7] Hamdan, M. A. (1972) On a characterization by conditional expectations. *Technometrics* **14** 497–499.
[8] Karlin, S. and Taylor, H. M. (1975) *A Second Course in Stochastic Processes*, Academic Press, London.
[9] Kestelman, H. (1959) *Modern Theories of Integration*, Dover, New York.
[10] Kotlarski, I. (1973) On a characterization of some distributions by conditional expectations. *Sankhya A* **34** 461–467.
[11] Kotlarski, I. and Hinds, W. E. (1975) Characterization theorems using conditional expectations. *J. Appl. Probab.* **12** 400–406.
[12] Kotz, S. and Shanbhag, D. N. (1980) Some new approaches to probability distributions. *Adv. Appl. Probab.* **12** 903–921.
[13] Kupka, J. and Loo, S. (1989) The hazard and vitality measures of ageing. *J. Appl. Probab.* **26** 532–542.
[14] Lebesgue, H. (1973) *L'Intégration*, Chelsea, New York.
[15] Lillo, R. E. and Martín, M. (1997) Characterization of Bayes estimators of the uniform scale parameters. *Statistics and Decisions* **17** 31–47.
[16] Meilijson, I. (1972) Limiting properties of the mean residual lifetime function. *Ann. Math. Statist.* **43** 354–357.
[17] Nagaraja, H. N. and Nevzorov, V. B. (1997) On characterizations based on record values and order statistics. *J. Statist. Plan. Inf.* **63** 271–284.
[18] Patil, G. P., Kotz, S. and Ord, J. K. (1975) *Statistical Distribution in Scientific Work*, Vol. 3, Reidel, Dordrecht, Holland, pp. 221–230.
[19] 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.
[20] Shanbhag, D. N. (1970) The characterizations for exponential and geometric distributions. *J. Amer. Statist. Assoc.* **65** 1256–1259.
[21] Watanabe, S. (1964) On discontinuous additive functionals and Lévy measures of a Markov process. *Japan J. Math.* **34** 53–70.

https://doi.org/10.1017/S0963548301004825 Published online by Cambridge University Press