EXPONENTIAL DISTRIBUTIONS ON PARTIALLY ORDERED ABELIAN SEMIGROUPS

by

Donald St. P. Richards
University of North Carolina at Chapel Hill

Abstract

Exponential random variables are defined on partially ordered Abelian semigroups by the condition that their "survival functions" are completely monotone. Several classical characterisations based on lack of memory properties, constancy of conditional expectations, order statistics and constant hazard rates are appropriately extended. The results are heavily dependent on the article of Berg et al (*Math. Ann.*, 223 (1976), 253-272).

Key words: Exponential distributions, Abelian semigroups, completely monotone, characterisations, survival function, lack of memory, conditional expectation, hazard rate, convolution semigroup.

AMS 1980 subject classification: 62E10, 60E05, 45A05.

December, 1981

1. Introduction

Recently, several articles devoted to characterisations of the exponential distribution have appeared; among others, we mention Davies (1981), Grosswald and Kotz (1981), Huang (1981), and Lau and Rao (1981). Together with the excellent monograph of Galambos and Kotz (1978), these serve to underline the importance and complexity of the exponential distribution.

In this article, we propose a generalisation of the exponential distribution to more abstract settings. The immediate motivation for this arises from the work of Puri and Rubin (1974), and Richards (1981). With $\mathbb{R}_+ = [0,\infty)$, the former authors show that the class of $\mathbb{R}_+^n$ - valued random vectors which have constant hazard rate are precisely those whose probability density functions are n-dimensional Laplace transforms of certain non-negative measures, i.e. the density functions are completely monotone on $\mathbb{R}_+^n$. The answer to a similar problem involving $\mathbb{R}_+^{n\times n}$, the cone of n x n positive definite matrices, was shown by Richards (1981) to have an analogous answer, viz. that certain density functions are completely monotone on $\mathbb{R}_+^{n\times n}$.

Both $\mathbb{R}_+^n$ and $\mathbb{R}_+^{n\times n}$ are examples of Abelian, partially ordered semigroups, and this is an indication of what may be a reasonable generalisation. One more clue is that in both examples mentioned above, the "survival functions" (defined with respect to the appropriate partial orders) of the vectors or matrices which possess constant hazard rate are also completely monotone. (Notice also that in Galambos and Kotz (1978), many results are proven by working with survival functions.)

Using the theory of completely monotone functions on Abelian semigroups due to Choquet (1954) and Berg, Christensen and Ressel (1976), we define exponential random variables on partially ordered semigroups by the requirement that their survival functions be completely monotone. Several classical characterisations based on lack of memory properties, constancy of conditional expectations, order statistics and constant hazard rates are appropriately extended.

As in some other articles, the central point of our treatment is the solution (in Section 2) of an integrated form of the Cauchy functional equation; from this follows many of the applications given in Section 3.

Berg (1975) has discussed "Gauss measures" on locally compact Abelian groups; we show in Section 4 that certain exponential random variables give rise to semigroups of measures which behave much the same as Gauss measures. Finally, Section 5 discusses concrete examples, and it suffices to say that if the semigroups can be embedded in finite dimensional spaces, then the exponential distributions are fairly well-behaved.

2. Notation and Preliminary Results

Let S be an Abelian semigroup, written additively. We assume, without loss of generality, that S contains an identity element, (denoted by 0); if not, we may adjoin an identity to S, and extend the addition to $S \cup \{0\}$ in the usual way.

Definition 2.1. A *character* on S is a function $\phi: S \to [0,1]$ such that

(a) $\phi(0) = 1$;

(b) $\phi(s + t) = \phi(s)\,\phi(t)$ for all $s, t \in S$.

By $\hat{S}$, we denote the set of all characters on S. It is clear that $\hat{S}$ is itself an Abelian semigroup with identity under the operation of pointwise multiplication. We shall later need to know that $\hat{S}$ is compact when endowed with the topology of pointwise convergence.

Definition 2.2. A function $f:S \to \mathbb{R}$ is *completely monotone* if and only if with $\Delta_t f(s) \equiv f(s + t) - f(s)$, $s, t \in S$, there holds

$$(-1)^n \Delta_{s_1} \Delta_{s_2} \dots \Delta_{s_n} f(s) \geq 0,$$

for any $n = 0,1,2, \dots$, and $s, s_1, s_2, \dots, s_n \in S$.

The convex cone M of completely monotone functions on S was introduced by Choquet (1954) (cf. Berg, Christensen and Ressel (1976)) in a study of capacities. A fundamental result established in both articles is the analogue of the classical Bernstein theorem for $\mathbb{R}_+$.

Theorem 2.1. *For any* $f \in M$, *there exists a unique Radon measure* ν_f *on* $\hat{S}$ *such that*

$$f(s) = \int_{\hat{S}} \phi(s) d\nu_f(\phi), \quad s \in S. \tag{2.1}$$

The correspondence $\nu_f \to f$ is called the *Laplace transform* of ν_f. We shall refer to ν_f as the *representing measure* for f.

We now assume that S is equipped with a partial order, $>$, which is compatible with the semigroup structure on S. More precisely, for all $s, t, u \in S$,

(a) $s > 0$ and $s > s$;

(b) $s + u > t + u$ whenever $s > t$;

(c) $s > t$, $t > u$ implies $s > u$.

An *initial* respectively, *terminal*) in S is a set of the form $\{t \in S: s > t\} \equiv \overleftarrow{s}$ (respectively, $\{t \in s: t > s\} \equiv \overrightarrow{s}$).

Let A be a σ- algebra of subsets of S which contains (or is generated by) the initials and terminals. Let (Ω, A, P) be a probability space, with the random variable $X:\Omega \to S$; we assume throughout that $P(X > s) \neq 0$ for all $s \in S$.

Definition 2.3 The random variable X has an *exponential distribution* if and only if the survival function $G(s) = P(X > s)$, $s \in S$, is completely monotone.

Remarks

1. Since $G(0) = 1$, then the representing measure for G is a probability measure on $\hat{S}$. Thus, when $S = \mathbb{R}_+$, the classical exponential distribution is an exponential in our sense. On $\mathbb{R}_+^{n\times n}$, the Wishart matrix on n+1 degrees of freedom is also an exponential, while on $\mathbb{R}_+^n$, the Marshall-Olkin exponential distributions are not.

2. In general, the probability distribution of X is not uniquely determined by the survival function G; exceptions to this are $\mathbb{R}_+^n$ and $\mathbb{R}_+^{n\times n}$ (cf. Simons (1974)).

3. Alternatively, we could define exponential random variables by requiring G to be a non-negative *positive definite* function (Berg et al (1976)). Note however, that G is necessarily monotonic decreasing on S, i.e. $G(s) \le G(t)$ whenever $s > t$. We conjecture that any function with these two properties is completely monotone, a result which we can prove for $\mathbb{Z}_+$ or for any Cartesian product of finitely many cyclic semigroups.

Let us now endow the *set* S with a topology under which S is locally compact (Hausdorff), and in which the σ-algebra A is the class of Borel sets. Let μ be a Borel measure on S; we assume that the support of μ is S, since otherwise, we may replace S by the smallest sub-semigroup containing the support of μ.

Suppose $f \in L^1(\mu)$ is bounded and non-negative, with

$$f(s) = \int_S f(s + t)\, d\mu\,(t) \quad , \; s \in S. \tag{2.2}$$

As in Lau and Rao (1981), we call this integral equation the *Integrated Cauchy Functional Equation* (ICFE).

For any $f \in L^1(\mu)$, let $\langle f,\mu\rangle \equiv \int_S f(s)d\mu\,(s)$, and $\hat{S}(\mu) = \{\phi \in \hat{S} \cap L^1(\mu): \langle \phi, \mu \rangle = 1\}$. Our method of solving (2.2) closely follows the proof of Theorem 2.3 in Berg et al (1976); it uses the Choquet integral representation theory. (See also the introduction in Furstenberg (1965).)

Theorem 2.2. *The function* f *satisfies* (2.2) *if and only if it is completely monotone, with the representing measure being supported by* $\hat{S}(\mu)$.

Proof. If $f \in M$ with ν_f supported by $\hat{S}(\mu)$, then Theorem 2.1 and an application of Fubini's theorem shows that f satisfies (2.2).

For the converse, let C be the set of all (bounded, non-negative) solutions to (2.2). Then, C is a proper, pointed, convex cone in the vector space $\mathbb{R}^S$ of all real-valued functions on S. Since C consists of bounded functions, then it is closed in the weak topology induced by $L^\infty(\mu)$. We let $C_1 = \{f \in C: f(0) \leq 1\}$; C_1 is a *cap* (cf. Meyer (1966)) of C and is compact and convex. Therefore, the non-zero extreme points of C_1 are the points of $C_2 = \{f \in C_1: f(0) = 1\}$ which are located on the extremal rays of C.

To determine these "extremal" points, observe that C is closed under translations, i.e. $f \in C$ if and only if $f_t \in C$ for all $t \in S$, $f_t(s) = f(s + t)$. Moreover, the ICFE implies (via approximation by simple functions) that f is a linear combination of its translates. Thus, $f \in C$ is extremal if and only if it is proportional to its translates f_t for all t in the support of μ, which shows that $f \in C \cap \hat{S} = \hat{S}(\mu)$.

It can also be proven that every $f \in \hat{S}(\mu)$ is extremal using the argument in the second part of the proof of Theorem 2.3 in Berg et al (1976). To conclude, we apply the Choquet theory, to represent each $f \in C$ as an integral over the extreme points of C_2. □

It seems difficult to deduce directly from (2.2) that f is completely monotone, without first solving the ICFE.

3. Applications to Characterisation Problems

3.1 The lack of memory properties.

In our setup, the (weak) lack of memory property becomes almost trivial.

Theorem 3.1. *If* $P(X > s + t \mid X > s) = P(X > t)$, *for all* $s, t \in S$, *then* X *has an exponential distribution.*

Proof. The hypothesis is that $G(s) = P(X > s)$ is a character, which is then completely monotone. □

Let Y be an S-valued random variable having probability distribution Q. We say that X has the *strong lack of memory* property if $P(X > Y + s \mid X > Y) = P(X > s)$, for all $s \in S$, and $P(X > Y) \neq 0$.

Theorem 3.2. *If* X *has strong lack of memory, and* Q *is supported by* S, *then* X *is exponentially distributed.*

Proof. Let $c = P(X > Y)$; then for any $s \in S$,

$$P(X > Y + s \mid X > Y) = c^{-1} \int_S \int_{\overrightarrow{t+s}} dP(u)\,dQ(t)$$

$$= \int_S G(s + t)\,dQ_1(t), \quad Q_1 = c^{-1}Q,$$

$$= G(s),$$

the last equality being the hypothesis. Since Q_1 is supported by S, then the result follows from Theorem 2.2 □

3.2 Characterisations via conditional expectations.

When $S = \mathbb{R}_+$, it is known (cf Klebanov (1980), Lau and Rao (1981)) that the criterion

$$E(h(X-s) \mid X > s) = E(h(X)), \text{ for all } s \in S, \tag{3.2}$$

implies that X has an exponential distribution, for a large class of functions $h : S \to \mathbb{R}_+$.

In the setting of semigroups however, we have no operation of subtraction. So, it is necessary to assume that there exists a group U, partially ordered under $>$, such that $S = \{s \in U : s > 0\}$. This assumption is maintained when we consider order statistics below.

Without loss of generality, we take $E(h(X)) = 1$ in (3.2). Further, we only consider h which satisfy the property that there exists a non-negative measure supported by S, also denoted by h, such that

$$h(s) = \int_{\overleftarrow{s}} dh(t)\ , \ s \in S. \tag{3.3}$$

We can now extend the classical results.

Theorem 3.3. *The criterion* (3.2) *holds if and only if* X *is exponentially distributed, in which case the representing measure for the survival function* G *is concentrated on* $\hat{S}(h)$.

Proof. Using (3.3), we obtain

$$E(h(X - s) \mid X > s) = \int_{\overrightarrow{s}} h(t - s)\, dP(t)/G(s)$$

$$= \int_{t\overleftarrow{-}s} \int_{\overrightarrow{s}} dP(t)\, dh(u)/G(s), \tag{3.4}$$

the last equality following from Fubini's theorem. Since U is an ordered group, the "inequalities" $t > s > 0$ and $t - s > u > 0$ are equivalent to $t > s + u > s > 0$, so that (3.2) is equivalent to

$$G(s) = \int_S G(s + t)\, dh(t) \quad , s \in S.$$

The conclusion follows from Theorem 2.2. □

A more difficult problem is the following; for any $s \in S$, define the random variable

$$(X - s)_+ = \begin{cases} X - s, & X > s; \\ 0 & , \text{otherwise}. \end{cases}$$

When $S = \mathbb{R}_+$ and $h : S \to \mathbb{R}_+$ is the identity function, Galambos and Kotz (1978, Theorem 2.3.3.) show that if the distribution of X is continuous at zero, then

$$E[h((X - s)_+)]E[h((X - t)_+)] \propto E[h((X - s - t)_+)], \tag{3.5}$$

for all $s, t \in S$, holds only when X is exponentially distributed. For an arbitrary semigroup S with any h satisfying (3.3) and $E(h(X)) = 1$, (3.5) is equivalent to

$$\int_S G(s + t)\, dh(t) = \phi(s), \text{ for all } s \in S, \tag{3.6}$$

where $\phi \in \hat{S}$. If X is exponentially distributed, then the uniqueness of Laplace transforms, Theorem 2.1 and (3.6) implies that $G \in \hat{S}(h)$. For general semigroups, however, we have not been able to prove a converse. In the cases $S = \mathbb{R}_+^n$ or $\mathbb{R}_+^{n\times n}$, a converse result can be obtained when $dh(t)$ is Lebesgue measure; the key step is that the characters are of the form $s \to \exp(- < s, t >)$ (on $\mathbb{R}_+^n$) and $s \to \exp(- \operatorname{trace}(st))$ (on $\mathbb{R}_+^{n\times n}$).

3.3 Characterisations via order statistics.

To speak of order statistics, we have to assume that S and U are linearly ordered. This effectively reduces the problems to the real line case so that there is nothing to be gained here. An alternative to the order statistics approach is to rephrase the hypotheses in terms of conditional distributions. As an example, we give a result below which in a sense extends a characterisation due to Puri and Rubin (1970) (cf. Galambos and Kotz (1978), Theorem 3.3.1.).

Theorem 3.4. *Let X and Y be i.i.d S-valued random variables. If the conditional distribution of X - Y given that X > Y is the same as the distribution of X, then X and Y are exponentially distributed.*

Proof. It suffices to show that X has the strong lack of memory property. But for any $s \in S$,

$$\begin{aligned} P(X > s) &= P(X - Y > s \mid X > Y) \\ &= P(X > Y + s \mid X > Y) \quad \square. \end{aligned}$$

3.4 Constant hazard rates.

Let ω be a non-negative measure on S which is invariant under translations. If the distribution of X is absolutely continuous with respect to ω and the corresponding density function is f, we say that X has *constant hazard rate* if $f(s)/G(s)$ is constant on S.

Theorem 3.5. *If X has constant hazard rate, then X is exponentially distributed.*

Proof. Since X has constant hazard rate, then there is a constant $c > 0$ such that for all $s \in S$,

$$f(s) = c \int_S f(t)\, d\omega(t)$$

$$= c \int_S f(s + t)\, d\omega(t),$$

the last following from ω being translation invariant. From Theorem 2.1, we have

$$f(s) = \int_{\hat{S}(c\omega)} \phi(s)\, d\nu(\phi) \ , \ s \in S,$$

from which it follows that G is completely monotone on S. □

4. Exponential Distributions and Gauss Measures

This section points out some parallels between the Gauss measures of Berg (1975), and certain convolution semigroups of probability measures determined by a particular type of exponential distribution.

Definition 4.1. (Berg et al (1976)). Let $f \in M_1 = \{g \in M : g(0) = 1\}$. Then, f is *infinitely divisible* if and only if for any $n \in \mathbb{N}$, there exists $f_n \in M_1$ such that $f(s) = (f_n(s))^n$ for all $s \in S$.

Let X be an exponential random variable with infinitely divisible survival function G. From Berg et al (1976), we know that any such G can be expressed as:

$$-\ell n G(s) = h(s) + \int_{\hat{S}_*} (1 - \phi(s))\, d\eta(\phi) \ , \ s \in S \tag{4.1}$$

Here, $h: S \to \mathbb{R}_+$ is linear on S, and η is a non-negative Radon measure on the locally compact monoid $\hat{S}_* = \hat{S} \setminus \{1\}$, $\{1\}$ denoting the identity character.

(4.1) is the *Lévy-Khintchine representation* associated with $-\ell nG$, and η is the Lévy measure for $-\ell nG$. We shall also need the result that G uniquely determines a *convolution semigroup* $(\eta_t)_{t>0}$ of probability measures on $\hat{S}$, such that as $t \downarrow 0$, $t^{-1}(\eta_t \mid \hat{S}_*) \to \eta$ vaguely.

Theorem 4.1. *Let the random variable* X *have infinitely divisible survival function* G. *Then, the following are equivalent:*

(i) $G \in \hat{S}$;

(ii) *The Lévy measure for* $-\ell nG$ *vanishes;*

(iii) *For any open neighbourhood* V *of* {1},

$$\lim_{t \downarrow 0} t^{-1}\eta_t(\hat{S} \setminus V) = 0. \tag{4.2}$$

The proof that (i) $\Longleftrightarrow$ (ii) follows from the discussion above and the uniqueness of the Lévy-Khintchine representation. Also, (ii) $\Longleftrightarrow$ (iii) may be established along the lines of Berg and Forst (1975), Theorem 18.27. We can also show that, as in the group case, if the convolution semigroup $(\eta_t)_{t>0}$ is of *local type* (Berg and Forst (1975)), then (4.2) holds.

It seems likely that this analogy between the Gauss measures and exponential distributions can be carried further.

5. Examples and Comments

In general, exponential distributions are not infinitely divisible (in the usual sense). One example is the nxn Wishart matrix on n+1 degrees of freedom; notice that here, the survival function is an infinitely divisible character. Even on finite semigroups, negative answers still arise (Martin-Löf (1969)).

Example. $S = \mathbb{R}_+^n$, with the usual vector addition. S is partially ordered by $s > t$ whenever t is "north-east" of s. As noted earlier, all characters are of the form $s \to \exp(- \langle s, t \rangle)$ for some $t \in S$. Thus, X is exponentially distributed whenever G(s) is the Laplace transform of a probability measure on $\mathbb{R}_+^n$; from this, we even know that G is continuous on $\mathbb{R}_+^n$.

For cones in $\mathbb{R}^n$, most of these remarks are valid, *mutatis mutandis*, so that the finite-dimensional exponential distributions are reasonably nice. On infinite-dimensional spaces however, it is possible to have very wild exponentials. Berg et al (1976, Section 8) show that on the semigroup $L_1^\infty([0,1])$, it is possible to have a representing measure which is concentrated on a set of characters all discontinuous at the origin. Another pathological example can be constructed using the semigroup $S = [0,1]$ as treated by Berg et al (1976), Section 7. (We would order S by : $s > t \iff s \leq t$.)

A final remark is that it would be good to have a definition of constant hazard rate which does not require the existence of a translation invariant measure on the semigroup S. (Recall that, in general, no such measure exists if the underlying space is infinite dimensional.) On $\mathbb{R}_+$, many authors define constant hazard rate by requiring that $-\ell n\, G$ is linear, but for general semigroups, this is far too restrictive.

Acknowledgements. This work was greatly enhanced by many discussions with I. Chakravarti, G. Kallianpur, V. Mandrekar, G. Simons and L. Stefanski. In particular, I am grateful to Professor Mandrekar for his innumerable suggestions and for bringing to my attention the article of Berg et al (1976), and for his help with the proof of Theorem 2.2.

REFERENCES

Berg, C. On the relation between Gaussian measures and convolution semigroups of local type. *Math. Scand.*, *37* (1975), 183-192.

Berg, C. and G. Forst. *Potential theory on locally compact Abelian groups.* Ergebnisse der Mathematik 87, Springer-Verlag, Berlin, Heidelberg, New York, (1975).

Berg, C., J.P.R. Christensen and P. Ressel. Positive definite functions on Abelian semigroups. *Math. Ann.* *223* (1976), 253-272.

Choquet, G. Theory of capacities. *Ann. Inst. Fourier, 5* (1954), 131-295.

Davies, L. A theorem of Deny with applications to characterization problems. *Lecture Notes in Math., 861* (1981), 35-41, (eds. Dugué, Lukacs and Rohatgi), Springer-Verlag, Berlin.

Galambos, J. and Kotz, S. *Characterizations of probability distributions. Lecture Notes in Math., 675,* Springer-Verlag, Berlin, 1978.

Grosswald, E. and Kotz, S. An integrated lack of memory characterization of the exponential distribution. *Ann. Inst. Statist. Math., 33* (1981), 205-211.

Huang, J.S. On a "lack of memory" property. *Ann. Inst. Statist. Math., 33* (1981).

Klebanov, L.B. On some characterizations of the exponential distributions. *Theory Prob. Appl., 25* (1980), 628-633.

Lau, K.-S. and Rao, C.R. Integrated Cauchy functional equation and characterizations of the exponential law. Tech. Rep. No. 81-10, 1981, Univ. of Pittsburgh.

Martin-Löf, P. Probability theory on discrete semigroups. *Z. Wahrschein. und verw. Gebiete, 4* (1965), 78-102.

Puri, P.S. and Rubin, H. A characterization based on the absolute difference of two i.i.d. random variables. *Ann. Math. Statist., 41* (1970), 2113-2122.

Puri, P.S. and Rubin. H. On a characterization of the family of distributions with constant multivariate failure rates. *Ann. Prob., 2* (1974), 738-740.

Richards, D. St. P. Hyperspherical models, fractional derivatives, and exponential distributions on matrix spaces. Preprint, 1981, Univ of the West Indies.

Simons, G. Generalized distribution functions: the linearly ordered case with applications to nonparametric statistics. *Ann. Prob., 2* (1974), 501-508.

Furstenberg, H. Translation-invariant cones of functions on semi-simple lie groups. *Bull. Amer. Math. Soc., 71* (1965), 271-326.

Meyer, P. A. *Probability and potentials*. Blaisdell, Waltham, 1966.