Applied Probability Trust (July 14, 1997)

NEW CHARACTERIZATIONS OF THE NO-AGING PROPERTY AND THE ℓ_1–ISOTROPIC MODEL

STEPHEN E. CHICK,* *University of Michigan*

MAX B. MENDEL,** *University of California*

Abstract

The no-aging property and the ℓ_1-isotropic model it implies have been introduced to overcome certain shortcomings of the exponential model. However, its definition is abstract and not very useful for practitioners. This paper presents several additional characterizations of the no-aging property. Included are (1) characterizations that appropriately generalize the memoryless property and the constant-failure-rate property of the exponential, (2) behavioral characterizations based on fair bets, and (3) geometric characterizations of the survival and density function and differential-geometric characterizations based on tensor methods.

NO-AGING, ℓ_1–ISOTROPIC MODEL, CHARACTERIZATION OF PROBABILITY MODELS, EXPONENTIAL MODEL, MEMORYLESS PROPERTY, HAZARD GRADIENT, BAYESIAN PROBABILITY, TENSOR METHODS IN PROBABILITY.

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 62E10
SECONDARY 60D05

* Postal address: Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109-2117, USA.

** Postal address: Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA.

1. Introduction

In this paper, we characterize the no-aging property and the joint probability models it implies. The no-aging property was proposed by Barlow and Mendel in [2]. They consider a finite, exchangeable population of lifetimes and argue that the correct probabilistic model for aging is given by a Schur-concave joint survival function (for the theory of Schur concavity, see e.g. Marshall and Olkin [12]). No-aging is described by a Schur-constant survival function and these functions are also called ℓ_1-isotropic. Their definition of no-aging is particularly appealing for a Bayesian who wishes to use the unconditional distribution of the lifetimes, rather than a distribution conditional on some parameter.

The problem with the no-aging property is that its definition in terms of Schur-constancy can be far removed from the type of assessments a practitioner is comfortable in making.

This paper compiles eight characterizations, five of them new, of the no-aging property and the ℓ_1-isotropic model in an attempt to make these concepts more useful to practitioners. The characterizations are compiled in the theorem in Section 2.

We have opted for characterizations that have a direct behavioral interpretation or an appealing geometric interpretation. The behavioral characterizations are particularly relevant to a Bayesian decision maker. They are stated in terms of fair bets. The geometric characterizations concern symmetries and invariances of the survival function and the density function. We also give several differential-geometric characterizations. The use of differential geometry is not very common in probability (see, however, Santalo [13]). At the same time, these characterizations are very compact and provide perhaps the best pictures.

Two of the characterizations are in the spirit of the classical constant-failure-rate and memoryless properties of the exponential. These two properties have played a central role in characterizing the exponential model; see, e.g., [4, 5, 6, 7, 8, 9, 10, 15]. These are useful for seeing how the no-aging property and the ℓ_1-isotropic model departs from the classical no-aging and the exponential model.

2. Characterizations

We state the various characterizations in the form of a theorem. An item-by-item discussion precedes the proof. We first introduce the main notation and we define the empirical failure frequency.

Consider a population

$$\vec{X} = (X_1, \ldots, X_N)$$

of lifetimes, i.e., a sequence of non-negative real-valued random variables. Let $\vec{X} \geq \vec{x}$ mean that $X_i \geq x_i$, for $i = 1, \ldots, N$. Write the joint survival function $\bar{F} : \mathbb{R}_+^N \to [0, 1]$ as

$$\bar{F}(\vec{x}) = \mathrm{Prob}(\vec{X} \geq \vec{x}), \tag{1}$$

and the density function $f : \mathbb{R}_+^N \to \mathbb{R}$ as

$$f(\vec{x}) = \frac{(-1)^N \partial^N \bar{F}(\vec{x})}{\partial x_1 \cdots \partial x_N},$$

when it exists.

Marginal survival and density functions of a single X_i or a subsequence $\vec{X}_n = (X_1, \ldots, X_n)$ will be denoted by the addition of a subscript to specify the arguments. For instance, denote

$$\bar{F}_{\vec{X}_n}(\vec{x}_n) = \mathrm{Prob}(X_1 \geq x_1, \ldots, X_n \geq x_n). \tag{2}$$

Survival probabilities conditional on the event A will be written $\bar{F}(\cdot \mid A)$.

Finally, let ∇ denote the gradient operator and call

$$-\nabla \log \bar{F}$$

the hazard gradient when it exists (see also Marshall [11]).

For the differential-geometric characterizations, let $\partial/\partial x_i$ be the infinitesimal vector in the i-th coordinate. The exterior derivative is denoted by d, the Lie derivative by $\pounds()$, the anti-symmetric product for tensors is denoted $\wedge$, and tensor saturation is denoted by $\lrcorner$. (See e.g. Burke [3]).

Definition: *For given lifetimes $\vec{x} = (x_1, \dots, x_N)$, the failure frequency is:*

$$\lambda_N = \frac{N}{\sum_{i=1}^{N} x_i},$$

where λ_N is written $\lambda_N(\vec{x})$ whenever the dependence on lifetimes is to be emphasized. When the lifetimes are unknown, we write $\Lambda_N(\vec{X}) = N / \sum_{i=1}^{N} X_i$.

We use the following definition of the no-aging property originally proposed by Barlow and Mendel [2].

Definition: *A population $\vec{X} = (X_1, \dots, X_N)$ has the no-aging property if $\bar{F}$ is a function of the failure frequency $\lambda_N(\vec{x})$ alone.*

The following theorem shows the equivalence of eight characterizations of the no-aging property, given certain differentiability assumptions. The characterizations make use of the fact that $\Lambda_N = \lambda_N$ defines a hyperplane in $\mathbb{R}_+^N$, just as $X_i = x_i$ defines a hyperplane. The definition appears as Char. 3. The equivalence of Char. 3 and Char. 4 is elementary, and the equivalence of Char. 3 and Char. 5 is due to Barlow and Mendel [2]. The other characterizations are new. Char. 1 can be viewed as a specialization of Spizzichino's characterization of the Schur-concave class [14].

A corollary describes the characterizations when the differentiability assumptions are not satisfied.

Theorem 1 Suppose that $f, d\bar{F}$ (as above) exist. The following are equivalent:

1. *$\bar{F}_{X_i}(x_i + h \mid \vec{X} \geq \vec{x}) = \bar{F}_{X_j}(x_j + h \mid \vec{X} \geq \vec{x})$ for all $\vec{x}$, and all $h > 0$.*

2. *The components of $-\nabla \log \bar{F}$ are identically-equal functions of λ_N alone, whenever $\bar{F} \neq 0$.*

3. *$\bar{F}$ is a function of λ_N alone (definition of the no-aging property).*

4. *f is a function of λ_N alone.*

5. *$\bar{F}_{\vec{X}_n}(\vec{x}_n)$ is an ℓ_1-isotropic survival function, i.e, for all n, $0 < n < N$.*

$$\bar{F}_{\vec{X}_n}(x_1, \dots, x_n) = \int \left[1 - \frac{\lambda_N}{N} \sum_{i=1}^{n} x_i\right]^{N-1} P(d\lambda_N).$$

Let $W_{i,j} = \partial/\partial x_i - \partial/\partial x_j$ *and* $\mathbf{W}$ *the wedge product of any* $(N-1)$ *linearly independent* $W_{i,j}$*. Let* $\mathbf{p}$ *be a probability* N*-form field on* $\mathbb{R}_+^N$*, and* $\mathbf{p}_{|\lambda_N}$ *the conditional probability* $(N-1)$*-form field, conditional on* $\{\Lambda_N(\vec{X}) = \lambda_N\}$*.*

6. $-d\log\bar{F} = g(\lambda_N)d\lambda_N$*, with* $g : \mathbb{R}_+ \to \mathbb{R}_+$*, whenever* $\bar{F}(\vec{x}) \neq 0$*.*

7. $\pounds_{W_{i,j}}(\mathbf{p}) = 0$ *for all* i, j*.*

8. $\mathbf{p}_{|\lambda_N} \lrcorner \mathbf{W}$ *is a function of* λ_N *alone.*

Remarks:

Char.1 is the appropriate generalization of the memoryless property. From Char. 1 it follows that the lifetimes are exchangeable. However, Char. 1 does *not* imply that the X_i's are independent. If we assume independence *in addition* to Char. 1, we find that:

$$\bar{F}_{X_i}(x_i + h|\vec{X} > \vec{x}) = \bar{F}_{X_i}(h), \ \ \forall i,$$

which is the memoryless property that characterizes the exponential.

Char. 1 has a behavioral interpretation: 'Given two similar components which have not yet failed, one would bet the same amount for the same return on the event that either component fails during an additional increment of usage, regardless of their ages.'

Char.2 is the appropriate generalization of the constant-failure-rate property. If we assume independence in addition to Char. 2, we find that:

$$\frac{d}{dx_i}\log\bar{F}_{X_i}(x_i) = \text{constant}, \ \ \forall i,$$

which characterizes i.i.d. exponential lifetimes (see Marshall [11]).

Char. 6 addresses several problems with the the use of the failure gradient (see remarks below).

Char.3 is the definition of the no-aging property and the ℓ_1-isotropic model. This is Schur-constancy of $\bar{F}$: $\bar{F}$ can be moved anywhere along a simplex $\{\Lambda_N(\vec{X}) = \lambda_N\}$ without changing its value. (See further Barlow and Mendel [2]).

Char. 3 is physically covariant under smooth, monotone increasing changes in the scales or units in which lifetime is measured.

Char.4 states that the density is uniform on simplexes $\{\Lambda_N(\vec{X}) = \lambda_N\}$. This uniformity, however, depends on the scale used for $\vec{X}$. f is *not* physically covariant, as changes in scale involve a Jacobian factor. See Char. 7 for a physically covariant version of Char. 4.

Char.5 gives a de Finetti-type representation for the family of ℓ_1-isotropic survival functions. For any N, this family strictly includes the i.i.d. exponentials, although the integrand represents dependent variables. When $N \to \infty$, it converges to a mixture of i.i.d. exponentials with rate λ_∞, the limiting failure frequency:

$$\bar{F}_{\vec{X}_n}(x_1, \dots, x_n) = \int_{\lambda_\infty = 0}^{\infty} \prod_{i=1}^{n} e^{-\lambda_\infty x_i} \, P(d\lambda_\infty).$$

(See also Barlow and Mendel [2].)

Char.6 improves upon Char. 2 in two important ways. (see Figure 1.)

First, Char. 6 is physically covariant. There are no explicit references to $\vec{X}$.

Second, the gradient operator is replaced with the exterior-derivative operator d, representing the differential or total derivative of $\log \bar{F}$. Unlike the gradient operator, the exterior derivative does not depend on a metric or inner product. This dependence on a metric is a not-so-well-known defect of Marshall's [11] hazard gradient: a gradient give the *direction* of greatest increase, but without a natural metric there is no natural notion of direction. Our definition makes $-d \log \bar{F}$ a 1-form rather than a vector, so a name such as "hazard 1-form" or "hazard differential" would be more appropriate.

Char.7 states that the joint density is uniform on simplexes as in Char. 4, but now using differential N-forms, thereby eliminating the references to $\vec{X}$. The W_{ij}'s are the vectors that leave λ_N invariant. By requiring that the Lie derivative of $\mathbf{p}$ vanishes, we obtain the invariant density. (See Figure 2.)

Force is a 1-form, so "force of mortality" is appropriate from this perspective.

Char.8 states that the conditional density on a simplex is uniform with respect to the natural notion of hyper-area on the simplex. **W** gives a unit of hyper-area on the simplex. $\mathbf{p}_{|\lambda_N} \lrcorner \mathbf{W}$ is a scalar function that gives the (infinitesimal) amount of probability per unit of hyper-area. If this is constant on the simplex, the conditional density is uniform with respect to **W**, as in Figure 3.

This characterization also has an interpretation in terms of fair bets: 'Given the failure frequency, one would bet the same amount for the same return on any outcome for $\vec{X}$.'

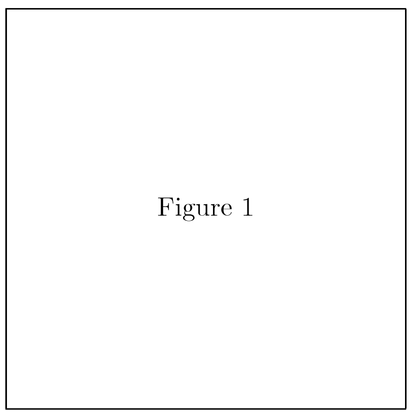


Figure 1. The hazard differentials at $\vec{x}$ and $\vec{x}'$; they are pictured by two infinitesimally-spaced level lines of the hazard function $-\log \bar{F}$ with the arrowhead flagging the positive direction (see also Burke [3]). Also pictured are the differentials of λ_N at $\vec{x}'$ and $\vec{x}''$. g is the factor with which $d\lambda_N$ is multiplied to obtain the hazard differential and this factor is constant on the simplex.

Proof. We show how Chars. 1 through 5 are equivalent to the definition in Char. 3. Chars. 6 and 7 are shown to be equivalent to their respective coordinate-based characterizations.

1.⇔3. Assume 3, $\bar{F} = \psi(\lambda_N)$ for some ψ. Then $\text{Prob}(X_i \geq x_i + h, \vec{X} \geq \vec{x}) = \text{Prob}(X_j \geq x_j + h, \vec{X} \geq \vec{x}) = \psi(N/(h + \sum_{i=1}^{N} X_i))$. 1 follows after dividing both sides of the first equality by $\bar{F}(\vec{x})$.

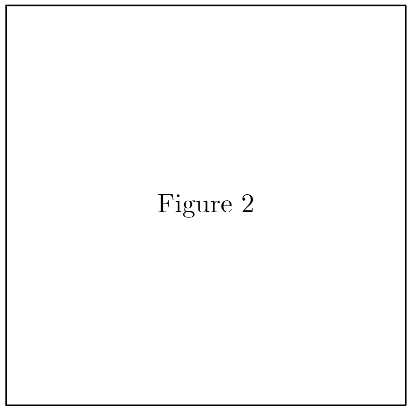


Figure 2. The probability N-form ($N = 2$) at two locations pictured by infinitesimal squares containing an infinitesimal unit of probability. Pulling $\mathbf{p}(\vec{x}')$ back via $W_{2,1}$ and subtracting it from $\mathbf{p}(\vec{x})$ shows that the Lie derivative vanishes.

Assume 1, multiply both sides by $\bar{F}(\vec{x})$, collect terms, and substitute $\vec{x} - he_j$ for $\vec{x}$. Thus $\mathrm{Prob}(\vec{X} \geq \vec{x} - he_j + he_i) = \mathrm{Prob}(\vec{X} \geq \vec{x})$ for all $0 < h < x_j$. Both $\bar{F}$ and λ_N are the same at $\vec{x} - he_j + he_i$ and $\vec{x}$. One can similarly shift lifetime between any pairs i, j while still conserving $\bar{F}$ and λ_N. As a consequence, $\bar{F}$ is a function of λ_N alone.

2.⇔3. Assume 2. Then $(\partial/\partial x_i - \partial/\partial x_j)\bar{F} = 0$, or $\nabla_{W_{i,j}}(\log \bar{F}) = 0$. This implies that $\log \bar{F}$ and hence $\bar{F}$ is a function of λ_N alone. If not, there would be point where $\nabla_{W_{i,j}} \log \bar{F} \neq 0$.

Assume 3. This implies $\log \bar{F}$ is a function of λ_N alone. Then

$$\begin{aligned}\nabla_{W_{i,j}}(\log \bar{F}) &= \frac{d\bar{F}}{d\lambda_N}\left(\frac{\partial \lambda_N}{\partial x_i} - \frac{\partial \lambda_N}{\partial x_j}\right) \\ &= 0.\end{aligned}$$

4.⇔3. Because the Lie brackets of the coordinate vectors in the $\vec{X}$ system all vanish, we have:

$$\frac{(-1)^N \partial^N}{\partial x_1 \cdots \partial x_N} \nabla_{W_{i,j}}(\bar{F}) = \nabla_{W_{i,j}}\left(\frac{(-1)^N \partial^N}{\partial x_1 \cdots \partial x_N} \bar{F}\right) = \nabla_{W_{i,j}}(f).$$

If $\bar{F}$ is a function of λ_N alone, then $\nabla_{W_{i,j}}(\bar{F}) = 0$. From the above equality

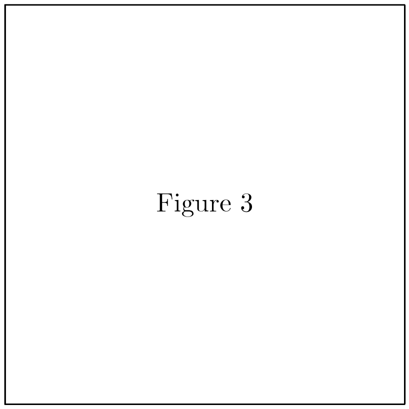


Figure 3. Here $N = 3$. The simplex contains the bivector $\mathbf{W}$ and the conditional probability 2-forms $\mathbf{p}_{|\lambda_N}$ on the simplex. $\mathbf{p}_{|\lambda_N} \lrcorner \mathbf{W}$ is the number of times the 2-form fits into the bivector (in the figure this is approximately 3 times) and this number is constant over the simplex as required.

it follows that $\nabla_{W_{i,j}}(f) = 0$, which implies 4. To prove $4 \to 3$ revert this argument.

5.⇔3. Assume 3. Then $\bar{F}(\cdot|\lambda_N)$ is constant on the simplex. Integrate out $x_{n+1}, \ldots x_N$. This is a Dirichlet integral and gives the result in 5 (see also, e.g., Barlow and Mendel [2]). That 5 implies 3 can be read off directly from the expression.

6.⇔2. Assume 6. Use the $\sharp$ operator for the Euclidean metric tensor (See e.g. Abraham and Marsden [1]). In $\vec{X}$ coordinates:

$$\sharp(-d\log\bar{F}) = -\frac{\partial \log \bar{F}}{\partial x_1}\frac{\partial}{\partial x_1} - \cdots - \frac{\partial \log \bar{F}}{\partial x_N}\frac{\partial}{\partial x_N} = -g(\lambda_N)\sharp(d\lambda_N).$$

But the components of $\sharp(d\lambda_N)$ depend only on λ_N and 2 follows.

Assume 2. Use the $\flat$ operator of the Euclidean metric tensor.

$$\flat(\nabla(-\log\bar{F})) = -\frac{\partial \log \bar{F}}{\partial x_1}dX_1 - \cdots - \frac{\partial \log \bar{F}}{\partial x_N}dX_N.$$

This is $-d\log\bar{F}$ in $\vec{X}$ coordinates. If the $\frac{\partial \log \bar{F}}{\partial x_i}$'s depend only on λ_N, 6 follows.

7.⇔4. Write the density N-form in $\vec{X}$ coordinates: $\mathbf{p} = f d\vec{X}$. The Leibniz property of the Lie derivative gives:

$$\pounds_{W_{i,j}}(f d\vec{X}) = \pounds_{W_{i,j}}(f) d\vec{X} + f \pounds_{W_{i,j}}(d\vec{X}).$$

But $\pounds_{W_{i,j}}(d\vec{X}) = 0$ and so we have $\pounds_{W_{i,j}}(\mathbf{p}) = 0$ if and only if $\pounds_{W_{i,j}}(f) = 0$, or $\nabla_{W_{i,j}}(f) = 0$, since for functions the Lie derivative is the directional derivative.

8.⇔7. Because $\pounds_{W_{i,j}}(h(\lambda_N)) = 0$ for any function h, 8 implies: $\pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N} \rfloor \mathbf{W}) = 0$ for all i, j. We have:

$$\pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N} \rfloor \mathbf{W}) = \pounds_{W_{i,j}}(\mathbf{W}) \rfloor \mathbf{p}_{|\lambda_N} + \mathbf{W} \rfloor \pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N}).$$

But $\pounds_{W_{i,j}}(\mathbf{W}) = 0$ and so we have 8 if and only if $\pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N}) = 0$ for all i, j. Now write $\mathbf{p} = \mathbf{p}_{\Lambda_N} \wedge \mathbf{p}_{|\lambda_N}$; $\mathbf{p}_{\Lambda_N} = P(d\lambda_N)$ is a marginal probability 1-form for Λ_N. We have

$$\pounds_{W_{i,j}}(\mathbf{p}_{\Lambda_N} \wedge \mathbf{p}_{|\lambda_N}) = \pounds_{W_{i,j}}(\mathbf{p}_{\Lambda_N}) \wedge \mathbf{p}_{|\lambda_N} + \mathbf{p}_{\Lambda_N} \wedge \pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N}).$$

But $\pounds_{W_{i,j}}(\mathbf{p}_{\Lambda_N}) = 0$ and so we have 7 also if and only if $\pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N}) = 0$ for all i, j. Now the $W_{i,j}$ have integral manifolds which are the level sets of λ_N and the equivalence follows.

Note that if $\text{Prob}(\Lambda_N < \mathcal{M}) = 1$ for some finite $\mathcal{M}$, then the support of the X_i is bounded on $(0, N\mathcal{M})$. If no such $\mathcal{M}$ exists, the support of the X_i is unbounded from above.

The theorem assumed the existence of the joint density f and the differential $d\bar{F}$ of the joint survival function. This condition might not hold, for example, when the mixing distribution in Char 5 is not absolutely continuous with respect to Lebesgue measure. Consider the characterizations when the differentiability assumptions are dropped. Char. 2 and Char. 6 become meaningless, since they fundamentally rely on the existence of $d\bar{F}$. Still, Char. 1, 3, 5, 8, and modified versions of 4 and 7 can be shown to be equivalent, as in the following corollary.

Corollary 1 When the existence of $f, d\bar{F}$ is not assumed, then Char. 1, 3, 5, 8 of the theorem and Char. 4′, 7′ are equivalent, where:

4′ *The conditional density $f_{\vec{X}_{N-1}|\lambda_N}$ exists and is a function of λ_N alone.*

7′ $\mathbf{p}_{|\lambda_N}$ *exists, and $\pounds_{W_{i,j}}(\mathbf{p}_{|\lambda_N}) = 0$ for all i, j.*

Proof. The equivalence of Char. 1, 3, and 5 established above did not make use of the differentiability conditions. The equivalence of Char. 4′, 7′ and 8 is almost identical to the analagous proofs above, except that the N-form $d\lambda_N \wedge \mathbf{p}_{|\lambda_N}$ should be used instead of the joint probability $\mathbf{p}$. Note that the proof only works when the conditional density $f_{\vec{X}_{N-1}|\lambda_N}$ exists. The equivalence of Char. 3 and 4′ similarly follows the analagous proof from the theorem above, except that $N-1$ directional derivatives along the $W_{i,j}$ are taken, rather than N coordinate-direction derivatives.

Acknowledgement: The authors appreciate the helpful comments of an anonymous referee.

References

[1] Abraham, R. and Marsden, J. E. (1985) *Foundations of Mechanics*. 2nd edn. Addison-Wesley, Redwood City, CA.

[2] Barlow, R. E. and Mendel, M. B. (1992) De Finetti-type representations for life distributions. *Journal of the American Statistical Association* **87,** 1116–1122.

[3] Burke, W. L. (1985) *Applied Differential Geometry*. Cambridge University Press, New York.

[4] Fosam, E. B. and Shanbhag, D. N. (1994) Certain characterizations of exponential and geometric distributions. *Journal of the Royal Statistical Society, Series B* **56,** 157–160.

[5] Galambos, J. and Hagwood, C. (1994) An unreliable server characterization of the exponential distribution. *Journal of Applied Probability* **31,** 274–279.

[6] Galambos, J. and Kotz, S. (1978) Characterizations of probability distributions. *Lecture Notes in Mathematics* **675,**.

[7] GATHER, U. (1989) On a characterization of the exponential distribution by properties of order statistics. *Statistics & Probability Letters* **7,** 93–96.

[8] HUANG, W. J. AND SHOUNG, J. M. (1993) Some characterization of the exponential and geometric distributions arising in a queueing system with an unreliable server. *Journal of Applied Probability* **30,** 985–990.

[9] KHALIL, Z. AND DIMITROV, B. (1994) The service time properties of an unreliable server characterize the exponential distribution. *Advances in Applied Probability* **26,** 172–182.

[10] LAU, K. S. AND RAO, C. R. (1982) Integrated Cauchy functional equation and chracterizations of the exponential law. *Sankhya A* **44,** 72–90.

[11] MARSHALL, A. (1975) Some comments on the hazard gradient. *Stochastic Processes and Their Applications* **3,** 293–300.

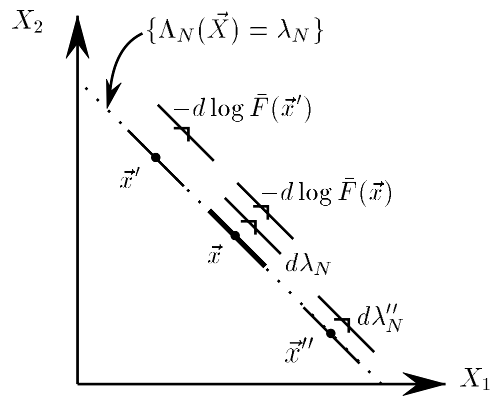
[12] MARSHALL, A. W. AND OLKIN, I. (1979) *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

[13] SANTALÓ, L. A. (1976) *Integral Geometry and Geometric Probability*. Addison-Wesley Publishing Co., Inc., Reading, MA.

[14] SPIZZICHINO, F. (1992) Reliability decision problems under conditions of aging. In *Bayesian Statistics 4*. ed. Bernardo., Berger., Dawid., and Smith. Oxford University Press, pp. 803–811.

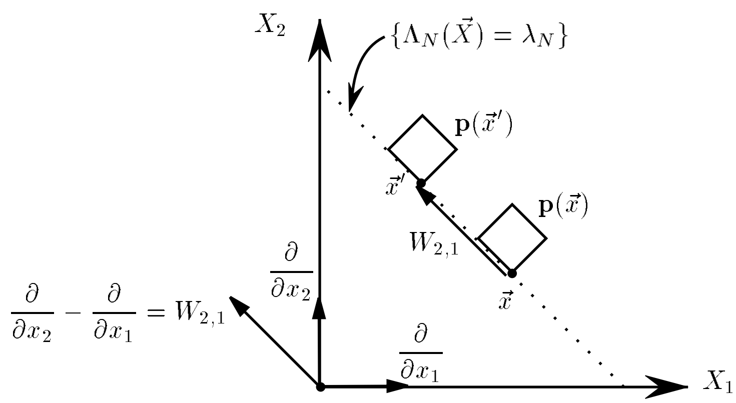
[15] TOO, Y. AND LIN, G. D. (1989) Characterizations of uniform and exponential distributions. *Statistics & Probability Letters* **7,** 357–359.

For Figure 1:

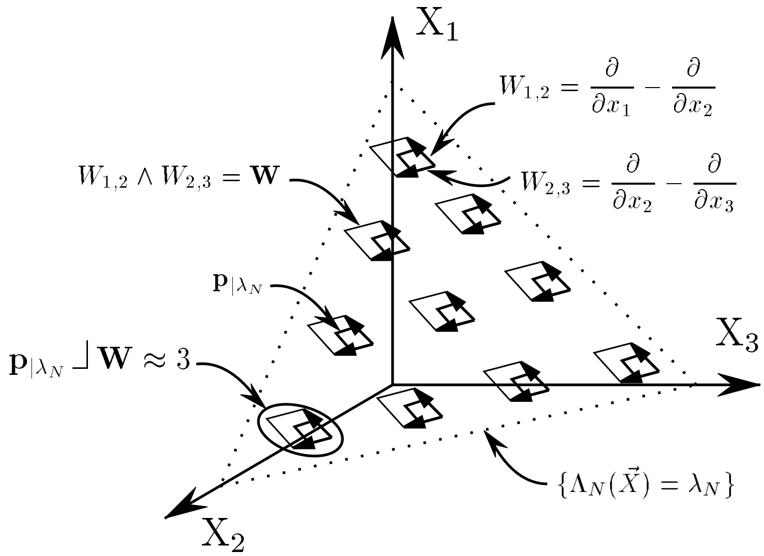


X_2
$\{\Lambda_N(\vec{X}) = \lambda_N\}$
$-d\log\bar{F}(\vec{x}')$
$\vec{x}'$
$-d\log\bar{F}(\vec{x})$
$\vec{x}$
$d\lambda_N$
$d\lambda_N''$
$\vec{x}''$
X_1

For Figure 2:


X_2
$\{\Lambda_N(\vec{X}) = \lambda_N\}$
$\mathbf{p}(\vec{x}')$
$\vec{x}'$
$\mathbf{p}(\vec{x})$
$W_{2,1}$
$\frac{\partial}{\partial x_2}$
$\vec{x}$
$\frac{\partial}{\partial x_2} - \frac{\partial}{\partial x_1} = W_{2,1}$
$\frac{\partial}{\partial x_1}$
X_1

For Figure 3:


X1
$W_{1,2} = \frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2}$
$W_{1,2} \wedge W_{2,3} = \mathbf{W}$
$W_{2,3} = \frac{\partial}{\partial x_2} - \frac{\partial}{\partial x_3}$
$\mathbf{p}_{|\lambda_N}$
X3
$\mathbf{p}_{|\lambda_N} \lrcorner \mathbf{W} \approx 3$
$\{\Lambda_N(\vec{X}) = \lambda_N\}$
X2