Estimation method of multivariate Exponential probabilities based on a Simplex Coordinates Transform

Niels J. Olieman †

Wageningen University, the Netherlands.

Bram van Putten ‡

Wageningen University, the Netherlands.

Abstract. A novel unbiased estimator, for estimating the probability mass of a multivariate exponential distribution over a measurable set, is introduced and is called the Exponential Simplex (ES) estimator. For any measurable set, the standard error of the ES-estimator is at most the standard error of the well known Monte Carlo (MC) estimator. For non-radially shaped measurable sets, the ES-estimator has a strictly smaller standard error than the MC-estimator. For ray-convex sets, such as convex sets, the ES-estimator can be expressed in a simple analytical form.

Keywords: Exponential Probability Estimation; Simplex Transformation; Monte Carlo Estimation; standard error efficiency; ray-convex sets; radially shaped sets

1. Introduction

In general, there are no analytical solutions for the computation of the probability mass of a multivariate exponential distribution over an arbitrary set. (See for instance Anderson, 1984; Narayan, 1996). This paper deals with a novel technique for estimating such a probability. The inspiration for the concept originates from i) the works of István Deák (1986), Alan Genz (1993) and Paul Somerville (1998) who studied integration of multivariate normal distributions and ii) practical experience of the first author with probability estimation methods in a game theoretic context (illustrated in Olieman and Hendrix, 2005; Dellink et al., 2005).

The following notation is used to express the probability of interest. The random vector $\underline{v}$ (NB: stochastic variates are underlined) with non-negative realisations $\mathbf{v} = [v_1, v_2, \dots, v_N] \in \mathbb{R}_+^N$ has independent Exponentially distributed elements $\underline{v}_i \sim \text{Exp}(1)$. It is well-known that $E(\underline{v}_i) = \text{var}(\underline{v}_i) = 1$. The Probability Density Function (PDF) of $\underline{v}$ is the function $f : \mathbb{R}^N \rightarrow \mathbb{R}_+$ with

$$f(\mathbf{v}) = \begin{cases} \prod_{i=1}^N \exp(-v_i) = \exp\left(-\sum_{i=1}^N v_i\right) & \text{for } \mathbf{v} \in \mathbb{R}_+^N \\ 0 & \text{elsewhere} \end{cases} \quad (1.1)$$

†Address for correspondence: Niels J. Olieman, Wageningen University, Department of Social Sciences, Operations Research and Logistics Group, P.O. Box 8130, 6700 EW Wageningen, the Netherlands.

E-mail: Niels.Olieman@wur.nl

‡Address for correspondence: Bram van Putten, Wageningen University, Biometris, P.O. Box 100, 6700 AC Wageningen, the Netherlands.

E-mail: Bram.vanPutten@wur.nl

The purpose of the estimation technique is to estimate

$$p = \Pr \{ \underline{v} \in \mathbb{A} \} = \int_{\mathbb{A}} f(\underline{v}) d\underline{v} \quad (1.2)$$

where $\mathbb{A}$ is an arbitrary measurable set in $\mathbb{R}^N$. Observe that an iso-probability density contour of function f , is a *positive orthant simplex* defined by $\sum_{i=1}^N v_i = r$ with $r \geq 0$ and $\underline{v} = [v_1, v_2, \dots, v_N] \in \mathbb{R}_+^N$. This feature will be exploited by the estimation method.

2. The Simplex Coordinates Transform

The integral expression in 1.2 does not have an analytical solution in general. As an analogue of the well-known polar coordinates transformation, in which the unit circle plays a central role, a "Simplex Coordinates Transformation" is defined, in which the unit simplex takes the role of the unit circle in the positive orthant. The transformation induces an equivalent expression for 1.2, which holds useful properties for estimation, which will be treated in more detail in section 3. The Simplex Coordinates Transformation (SCT) is defined by the following mapping

$$T : \mathbb{Q} \longrightarrow \mathbb{R}_+^N \text{ with} \quad (2.1)$$

$$T(r, s_1, s_2, \dots, s_{N-1}) = \left(rs_1, rs_2, \dots, rs_{N-1}, r(1 - \sum_{i=1}^{N-1} s_i) \right) \quad (2.2)$$

where $\mathbb{Q} = (\mathbb{R}_+, \mathbb{S}^{N-1})$ and $\mathbb{S}^{N-1} = \left\{ \underline{s} \in \mathbb{R}_+^{N-1} \mid \sum_{i=1}^{N-1} s_i \leq 1 \right\}$. Note that for $r = 1$, coordinates from the lower dimensional space $\mathbb{S}^{N-1}$ are mapped on unit simplex coordinates in $\mathbb{R}^N$. The variable $r \in \mathbb{R}_+$ can be interpreted as a magnification factor, similar to the radius variable in polar coordinates transformation.

Transformation T has the following properties:

- T is injective: $T(q_1) \neq T(q_2)$ if $q_1 \neq q_2$, for $q_1, q_2 \in \mathbb{Q}$
- T is surjective ("onto"): $\bigcup_{q \in \mathbb{Q}} T(q) = \mathbb{R}_+^N$
- The Jacobian $J : \mathbb{Q} \longrightarrow \mathbb{R}$ is

$$\det \frac{\delta T(q)}{\delta q} = \det \begin{bmatrix} s_1 & r & 0 & 0 & \dots & 0 \\ s_2 & 0 & r & 0 & \dots & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ s_{N-1} & 0 & 0 & 0 & \dots & r \\ 1 - s_1 - s_2 - \dots - s_{N-1} & -r & -r & -r & \dots & -r \end{bmatrix} = (-1)^N r^{N-1} \quad (2.3)$$

The latter step can be obtained by summing up the first $N - 1$ rows to the last row and by expanding the resulting determinant along its last row. Observe that J has constant sign with absolute value equal to r^{N-1} .

These properties imply that 1.2 can be rewritten as

$$\Pr \{\underline{v} \in \mathbb{A}\} = \int_{\mathbb{S}^{N-1}} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \mathbf{s}) \exp(-r) r^{N-1} dr d\mathbf{s} \quad (2.4)$$

where $I_{\mathbb{A}}^{SC} : \mathbb{Q} \rightarrow \{0, 1\}$ is the indicator function of $\mathbb{A}$, which expressed in Simplex Coordinates (SC) reads

$$I_{\mathbb{A}}^{SC}(r, \mathbf{s}) = \begin{cases} 1 & \text{if } T(r, \mathbf{s}) \in \mathbb{A} \\ 0 & \text{elsewhere} \end{cases} \quad (2.5)$$

and factor $\exp(-r)$ is obtained via

$$f(T(r, \mathbf{s})) = \exp\left(-\sum_{i=1}^{N-1} r s_i\right) \exp\left(-r\left(1 - \sum_{i=1}^{N-1} s_i\right)\right) = \exp(-r) \quad (2.6)$$

for $r \in \mathbb{R}_+$ and $\mathbf{s} \in \mathbb{S}^{N-1}$.

3. Exponential Simplex (ES) Estimation Method

The theory developed in the previous sections will be used to develop the so-called *Exponential Simplex* (ES) estimation method, which is a promising alternative to the well known Monte Carlo estimation method. In section 4 the efficiency of both methods is compared.

Let $\underline{\mathbf{s}}$ be a uniformly distributed random vector[§] with support set $\mathbb{S}^{N-1}$ for certain $N \in \mathbb{N}$. The estimation procedure is based on the following theorem

Theorem 1

$$E_{\underline{\mathbf{s}}} \left[\frac{1}{(N-1)!} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \underline{\mathbf{s}}) \exp(-r) r^{N-1} dr \right] = \Pr \{\underline{v} \in \mathbb{A}\} \quad (3.1)$$

Proof. From standard calculus it follows that the *volume* of the support of $\underline{\mathbf{s}}$ is $\int_{\mathbb{S}^{N-1}} 1 d\mathbf{s} = \frac{1}{(N-1)!}$, such that the PDF of $\underline{\mathbf{s}}$ is defined by a constant function over its support

$$g(\mathbf{s}) = \begin{cases} (N-1)! & \text{for } \mathbf{s} \in \mathbb{S}^{N-1} \\ 0 & \text{elsewhere} \end{cases}$$

[§]For the application of such a method, it is relevant to mention that methods for generating random points in a polytope (such as $\mathbb{S}^{N-1}$) are discussed in detail in Devroye (1986).

The left-hand side of expression 3.1 leads to integral 2.4 via

$$\begin{aligned}
& E_{\underline{s}} \left[\frac{1}{(N-1)!} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \underline{s}) \exp(-r) r^{N-1} dr \right] \\
&= \int_{\mathbb{S}^{N-1}} g(\underline{s}) \int_0^\infty \frac{1}{(N-1)!} I_{\mathbb{A}}^{SC}(r, \underline{s}) \exp(-r) r^{N-1} dr d\underline{s} \\
&= \int_{\mathbb{S}^{N-1}} \frac{(N-1)!}{(N-1)!} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \underline{s}) \exp(-r) r^{N-1} dr d\underline{s}
\end{aligned} \tag{3.2}$$

which is equivalent to probability $\Pr \{\underline{v} \in \mathbb{A}\}$ ■

Let $\underline{s}^{(1)}, \dots, \underline{s}^{(M)}$ be M independent copies of random vector $\underline{s}$. The random variable $\hat{\underline{p}}_{ES}^{[M]}$ is called the *ES-estimator* and is defined as

$$\hat{\underline{p}}_{ES}^{[M]} = \frac{1}{M} \sum_{m=1}^M \frac{1}{(N-1)!} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \underline{s}^{(m)}) \exp(-r) r^{N-1} dr \tag{3.3}$$

From theorem 1 it follows directly

Corollary 1

$\hat{\underline{p}}_{ES}^{[M]}$ is an unbiased estimator of $p = \Pr \{\underline{v} \in \mathbb{A}\}$.

In the following, for fixed $\underline{s} \in \mathbb{S}^{N-1}$, the set $\{T(r, \underline{s}) | r \in \mathbb{R}_+\}$ is called a *ray*.

Definition 1 The measurable set $\mathbb{A}$ is called ray-convex if $T(r_1, \underline{s}) \in \mathbb{A}$ and $T(r_2, \underline{s}) \in \mathbb{A}$ for $r_1, r_2 \geq 0$ and $\underline{s} \in \mathbb{S}^{N-1}$ implies that $T(\alpha r_1 + (1-\alpha)r_2, \underline{s}) \in \mathbb{A}$ for $0 \leq \alpha \leq 1$.

Note that a convex set is also ray-convex, but the converse is not true.

Definition 2 For fixed $\underline{s} \in \mathbb{S}^{N-1}$, let

$$\begin{aligned}
a(\underline{s}) &= \inf \{r \in \mathbb{R}_+ | T(r, \underline{s}) \in \mathbb{A}\} \\
b(\underline{s}) &= \sup \{r \in \mathbb{R}_+ | T(r, \underline{s}) \in \mathbb{A}\}
\end{aligned}$$

Remark that $b(\underline{s})$ can be ∞ .

Theorem 2 For a ray-convex set $\mathbb{A}$

$$\hat{\underline{p}}_{ES}^{[M]} = \frac{1}{M} \sum_{m=1}^M \left\{ f_N(a(\underline{s}^{(m)})) - f_N(b(\underline{s}^{(m)})) \right\} \tag{3.4}$$

where

$$f_N(x) = \exp(-x) \sum_{i=0}^{N-1} \frac{x^i}{i!} \tag{3.5}$$

Proof.

By ray-convexity, formula (3.3) can be written as

$$\hat{\underline{p}}_{ES}^{[M]} = \frac{1}{M} \sum_{m=1}^M \frac{1}{(N-1)!} \int_{a(\underline{s}^{(m)})}^{b(\underline{s}^{(m)})} \exp(-r) r^{N-1} dr \quad (3.6)$$

By induction

$$\frac{1}{(N-1)!} \int_a^b \exp(-r) r^{N-1} dr = \exp(-a) \sum_{i=0}^{N-1} \frac{a^i}{i!} - \exp(-b) \sum_{i=0}^{N-1} \frac{b^i}{i!} \quad (3.7)$$

■

In the situation the measurable set $\mathbb{A}$ is a finite union $\bigcup_{h=1}^H \mathbb{A}_h$ of disjunct ray-convex sets $\mathbb{A}_h$, the probability estimator of $\mathbb{A}$ is the sum of the ES-estimators of the individual sets.

4. Comparison of Exponential Simplex (ES) and Monte Carlo (MC) estimation methods

Up till now, *Monte Carlo* (MC) estimation is the standard methodology used for the estimation of 1.2 see Ghahramani (2000); Grimmett and Stirzaker (2001); Rice (1995) among others. Let $\underline{v}^{(1)}, \dots, \underline{v}^{(M)}$ be M independent copies of random vector $\underline{v}$. The random variable $\hat{\underline{p}}_{MC}^{[M]}$ is defined as

$$\hat{\underline{p}}_{MC}^{[M]} = \frac{1}{M} \sum_{m=1}^M I_{\mathbb{A}}^{CC}(\underline{v}^{(m)}) \quad (4.1)$$

where $I_{\mathbb{A}}^{CC} : \mathbb{R}^N \rightarrow \{0, 1\}$ is the indicator function in Cartesian Coordinates (CC), given by

$$I_{\mathbb{A}}^{CC}(\underline{v}) = \begin{cases} 1 & \text{if } \underline{v} \in \mathbb{A} \\ 0 & \text{else} \end{cases} \quad (4.2)$$

We will compare the Exponential Simplex (ES) estimation method and the Monte Carlo estimation method, with respect to their statistical quality, expressed in terms of standard error (se)

Theorem 3

$$se(\hat{\underline{p}}_{ES}^{[M]}) \leq se(\hat{\underline{p}}_{MC}^{[M]}) \quad (4.3)$$

Proof.

It is sufficient to prove the equation for $M = 1$. As $0 \leq (\hat{\underline{p}}_{ES}^{[1]})^2 \leq \hat{\underline{p}}_{ES}^{[1]} \leq 1$, it holds that

$$\begin{aligned} var(\hat{\underline{p}}_{ES}^{[1]}) &= E(\hat{\underline{p}}_{ES}^{[1]})^2 - (E\hat{\underline{p}}_{ES}^{[1]})^2 \\ &\leq E(\hat{\underline{p}}_{ES}^{[1]}) - (E\hat{\underline{p}}_{ES}^{[1]})^2 = p - p^2 = var(\hat{\underline{p}}_{MC}^{[1]}) \end{aligned}$$

For general M , from in-dependency and unbiasedness of both estimators, it follows now

$$se\left(\hat{\underline{p}}_{ES}^{[M]}\right) = \left[\frac{var(\hat{\underline{p}}_{ES}^{[1]})}{M}\right]^{1/2} \leq \left[\frac{var(\hat{\underline{p}}_{MC}^{[1]})}{M}\right]^{1/2} = se\left(\hat{\underline{p}}_{MC}^{[M]}\right) \quad (4.4)$$

■

Let $\mathbb{Q}_x^{\mathbb{A}} = \left\{ \mathbf{s} \in \mathbb{S}^{N-1} \left| \frac{1}{(N-1)!} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \mathbf{s}) \exp(-r) r^{N-1} dr = x \right. \right\}$ for $x \in [0, 1]$.

Definition 3 *The measurable set $\mathbb{A} \subseteq \mathbb{R}_+^N$ is called to have a radial shape if $\mathbb{Q}_0^{\mathbb{A}} \cup \mathbb{Q}_1^{\mathbb{A}}$ is dense in $\mathbb{S}^{N-1}$ (i.e. any ray corresponding $\mathbf{s} \in \mathbb{S}^{N-1}$ is either completely inside or completely outside $\mathbb{A}$).*

Theorem 4 *For a measurable set $\mathbb{A}$ it holds that*

$$se\left(\hat{\underline{p}}_{ES}^{[M]}\right) = se\left(\hat{\underline{p}}_{MC}^{[M]}\right)$$

if and only if $\mathbb{A}$ has a radial shape.

Proof.

In the proof of Theorem 3, the equality

$$E\left(\hat{\underline{p}}_{ES}^{[1]}\right)^2 = E\left(\hat{\underline{p}}_{MC}^{[1]}\right)^2$$

is equivalent to

$$E\left(\hat{\underline{p}}_{ES}^{[1]}\right)^2 = E\left(\hat{\underline{p}}_{ES}^{[1]}\right)^2$$

which holds if and only if all realisations of the estimator $\underline{p}_{ES}^{[1]}$ are either 0 or 1, i.e. if and only if the set

$$\left\{ \mathbf{s} \in \mathbb{S}^{N-1} \left| \frac{1}{(N-1)!} \int_0^\infty I_{\mathbb{A}}^{SC}(r, \mathbf{s}) \exp(-r) r^{N-1} dr \in \{0, 1\} \right. \right\}$$

is dense in $\mathbb{S}^{N-1}$, i.e. is equal to $\mathbb{S}^{N-1}$ apart from a set of measure zero. ■

From theorem 4 it follows directly

Corollary 2

For a measurable set $\mathbb{A}$ it holds that

$$se\left(\hat{\underline{p}}_{ES}^{[M]}\right) < se\left(\hat{\underline{p}}_{MC}^{[M]}\right)$$

if and only if $\mathbb{A}$ does not have a radial shape.

For a radially shaped set $\mathbb{A}$, the difference between the MC standard error and ES standard error is minimal (i.e. is zero). Conversely, the maximum standard error difference occurs, for instance, when $\mathbb{A}$ is a dense set in the positive orthant, bounded by a simplex. In this situation $a(\mathbf{s})$ and $b(\mathbf{s})$ are constant for any $\mathbf{s} \in \mathbb{S}^{N-1}$ and consequently all realisations of the ES-estimator equal $\Pr\{\mathbf{g} \in \mathbb{A}\}$. In this case the standard error of the ES-estimator is zero whereas the standard error of MC-estimator is $(\frac{1}{M}p(1-p))^{\frac{1}{2}}$ with maximum $\frac{1}{2\sqrt{M}}$ in case $p = \frac{1}{2}$.

5. Concluding Remarks

Apart from being superior to MC-estimator in the above mentioned statistical sense, a main advantage of ES-estimator is that, for ray-convex sets, it can be expressed in a very simple analytical form following Theorem 2. In case of a finite union of disjunct ray-convex sets, the ES-estimator takes the form of the sum of the corresponding analytical expressions. We remark that the class of (finite unions of disjunct) ray-convex sets is of much practical interest, as it contains the class of (finite unions of disjunct) convex sets. The above mentioned advantages of ES-estimation to MC-estimation could become less prominent in those practical cases in which expressions $a(\mathbf{s})$ and $b(\mathbf{s})$ given in Definition 2 are computationally complex. An intermediate solution in that case could be to approximate, conditionally on $\mathbf{s}$, the intersection points $a(\mathbf{s})$ and $b(\mathbf{s})$ following numerical methods.

References

- Anderson, T. W. (1984). *An introduction to multivariate statistical analysis 2ed.* New York: John Wiley & Sons.
- Deák, I. (1986). Computing probabilities of rectangles in case of multinormal distribution. *Journal of Statistical Computation and Simulation* 26, 101–114.
- Dellink, R., M. Finus, and N. J. Olieman (2005). Coalition formation under uncertainty: the stability likelihood of an international climate agreement. Technical Report Nota di Lavoro 98.2005, FEEM.
- Devroye, L. (1986). *Non-Uniform Random Variate Generation.* New York: Springer-Verlag.
- Genz, A. (1993). Comparison of methods for the computation of multivariate normal probabilities. *Computing Science and Statistics* 25, 400–405.
- Ghahramani, S. (2000). *Fundamentals of Probability Theory.* Upper Saddle River, New Jersey 07458: Prentice Hall.
- Grimmett, G. R. and D. R. Stirzaker (2001). *Probability and Random Processes 3ed.* Great Clarendon Street, Oxford ox2 6dp: Oxford University Press.
- Narayan, C. G. (1996). *Multivariate Statistical Analysis.* 270 Madison Avenue, New York, New York 10016: Marcel Dekker, INC.
- Olieman, N. J. and E. M. T. Hendrix (2005). Stability likelihood of coalitions in a two-stage cartelgame: An estimation method. *European Journal of Operational Research* (In Press; available via www.sciencedirect.com).

Rice, J. A. (1995). *Mathematical statistics and data analysis (2ed)*. Belmont, California 94002: Wadsworth Publishing Company.

Somerville, P. N. (1998). Numerical computation of multivariate normal and multivariate-t probabilities over convex regions. *Journal of Computational and Graphical Statistics* 7(4), 529–544.