Duality for real and multivariate exponential families

Gérard Letac\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a}Institut de Mathématiques de Toulouse, 118 route de Narbonne 31062 Toulouse, France.

Abstract
Consider a measure $\mu$ on $\mathbb{R}^n$ generating a natural exponential family $F(\mu)$ with variance function $V_F(\mu)(m)$ and Laplace transform
\[ \exp(\ell_\mu(s)) = \int_{\mathbb{R}^n} \exp(-\langle s, x \rangle) \mu(dx). \]
A dual measure $\mu^*$ satisfies $-\ell_{\mu^*}'(-\ell_{\mu}')(s) = s$. Such a dual measure does not always exist. One important property is $\ell_{\mu^*}''(m) = (V_F(\mu)(m))^{-1}$, leading to the notion of duality among exponential families (or rather among the extended notion of T exponential families $TF$ obtained by considering all translations of a given exponential family $F$).

Keywords: Dilogarithm distribution, Landau distribution, large deviations, quadratic and cubic real exponential families, Tweedie scale, Wishart distributions.

2020 MSC: Primary 62H05, Secondary 60E10

1. Introduction
One can be surprised by the explicit formulas that one gets from large deviations in one dimension. Consider iid random variables $X_1, \ldots, X_n, \ldots$ and the empirical mean $X_n = (X_1 + \cdots + X_n)/n$. Then the Cramér theorem [6] says that the limit
\[ \Pr(X_n > m)^{1/n} \xrightarrow{n \to \infty} \alpha(m) \]
does exist for $m > \mathbb{E}(X_1)$. For instance, the symmetric Bernoulli case $\Pr(X_n = \pm 1) = \frac{1}{2}$ leads, for $0 < m < 1$, to
\[ \alpha(m) = (1 + m)^{-1-m}(1 - m)^{-1+m}. \quad (1) \]
For the bilateral exponential distribution $X_n \sim \frac{1}{2} e^{-|x|} dx$ we get, for $m > 0$,
\[ \alpha(m) = \frac{1}{2} e^{-\sqrt{1+m^2}} (1 + \sqrt{1+m^2}). \]
What strange formulas! Other ones can be found in [17], Problem 410. The present paper is going to interpret in certain cases the function $m \mapsto 1/\alpha(m)$ as the Laplace transform of a certain dual measure $\mu^*$ which can be deduced explicitly from the distribution $\mu$ of $X_n$. Since the Cramér theorem can be seen as a result about one dimensional exponential families, we develop the idea in this framework, using the large box of examples obtained from the theory of variance functions of exponential families initiated by the article of Carl Morris [21] in 1982. An even simpler example of duality is provided by the Tweedie families (Barl cv and Enis [1], Jørgensen [9] and Tweedie [22]): the variance function $Am^p$ with $p > 1$ has dual $Bm^q$ with $\frac{1}{p} + \frac{1}{q} = 1$ for suitable pairs $(A, B)$ (see (8)). For instance, the Inverse Gaussian family $Am^3$ has dual $Bm^{3/2}$.

The Poisson distribution, with the one of the simplest variance functions $V_F(m) = m$, leads to the study of exponential family with variance function $e^m$ generated up to a translation by the unsymmetric stable law with Laplace
transform $e^{-x}s'$, also called Landau distribution. This gives tools for describing duals of other familiar exponential families. The cases of the normal and gamma families are very simple, being self dual, but other familiar cases like the negative binomial, the Bernoulli distribution and the cubic families are tougher. Finally, we consider another family, which is the dilogarithm family with variance function $e^{m} - 1$ as well as the one with variance function $\sinh m$. Like the normal and gamma families, they have the remarkable property to be self dual.

The definition of dual measures makes sense also in $\mathbb{R}^n$ while the probabilistic interpretation in terms of large deviations is lost. However we consider several cases in $\mathbb{R}^n$: the multinomial distribution, the Wishart ones and other quadratic families as classified by Casalis [21].

We proceed as follows: the notion of duality leads us unfortunately to change a bit the tradition about the exponential families. Indeed we will use $e^{-x}$ instead of $e^{x}$ in order to obtain later more readable formulas. This is explained in Section 2, together with the description of the classical objects attached to an exponential family.

In the preceding lines we have been vague about duality. Section 3 gives proper definitions, explaining what we call a dual measure $\mu'$ of $\mu$ and showing that some measures have no dual. We explain also what a $T$ exponential family $TF$ is. It is nothing but an exponential family $F$ plus all its translations. Indeed, talking about the dual $F^*$ of an exponential family $F$ does not exactly make sense, while the dual $TF^*$ of $TF$ does. Section 3 gives also the link with large deviations.

Section 4 concentrates on the $TF$ when the variance function $V_F(m)$ is $e^{m}$ and some parent distributions. It also give details on what we call Lévy measures of types 0, 1 and 2. Of course, large parts of this material are well known from probabilists and statisticians: exponential families, variance functions, Lévy measures, Landau distribution. It was necessary to expose them again for commodity of reading. This section contains crucial calculations for the sequel in Proposition 7.

Section 5 applies the results of Section 4 to the description of the duals of the Morris and the cubic families, with the surprising fact that they exist all with the only exception of the hyperbolic family with variance function $m^2 + 1$.

Section 6 describes the self dual dilogarithm distribution $\mu$ on the set $\mathbb{N} = \{0, 1, \ldots \}$ of integers defined by

$$\sum_{n=0}^{\infty} \mu(n)z^n = \exp\left(\sum_{k=1}^{\infty} \frac{1}{k^2}(z^k - 1)\right)$$

which, for $m > 0$, generates the exponential family with variance function $e^{m} - 1$. Since the consideration of this exponential family and of a set of parent distributions is not done in the literature, we develop some of their properties, somewhat deviating from the study of duality. For instance, if $N$ is the standard Gaussian distribution, then the variance of the exponential family generated by the convolution $N * \mu$ is $e^{m} + 1$.

Section 7 considers the $\mathbb{R}^n$ case: The multinomial distribution has a very explicit dual expressed in terms of the Landau distribution. The Wishart distribution is self dual as the one dimensional gamma distribution. We prove some negative results, like the fact that the multivariate negative binomial law has no dual.

Section 8 discusses open problems.

2. Laplace and bilateral Laplace transforms

At first the exponential families and Laplace transforms are considered. A certain tradition among statisticians (see Morris [21]) as opposed to physicists, and may be to probabilists, defines the Laplace transform of a positive measure $\mu$ on $\mathbb{R}$ and $\mathbb{R}^n$ as follows:

$$L_\mu(\theta) = \int_{\mathbb{R}} e^{\langle \theta, x \rangle} \mu(dx).$$

(2)

From the H"older inequality the set $D(\mu) = \{\theta; L_\mu(\theta) < \infty\}$ is a convex set, and the function $k_\mu = \log L_\mu$ is convex on $D(\mu)$. Actually $k_\mu$ is strictly convex outside of the particular case where $\mu$ is concentrated on one point in the case of $\mathbb{R}$, or on an affine hyperplane in the case of $\mathbb{R}^n$. To avoid trivialities one introduces the interior $\Theta(\mu)$ of $D(\mu)$. One calls $M(\mathbb{R}^n)$ the set of $\mu$ which are not concentrated on an affine hyperplane and such that $\Theta(\mu)$ is not empty. Such a $\mu$ generates a set of probabilities

$$P(\theta, \mu)(dx) = e^{\langle \theta, x \rangle - k_\mu(\theta)} \mu(dx)$$

2
and $F = F(\mu) = \{P(\theta, \mu) \; ; \theta \in \Theta(\mu)\}$ is called the natural exponential family generated by $\mu$. Note that $\mu$ is not necessarily bounded: simple examples on $\mathbb{R}$ like $\mu(dx) = 1/(\pi\sqrt{x})dx$ or $\sum_{n=0}^{\infty} \delta_n$ generate the important families of exponential distributions or geometric discrete laws. Omitting 'natural', we will say always exponential family for short. Objects linked to $F$ are the mean $m$ of $P(\theta, \mu)$, the domain of the means $M_F$ and the inverse function $\psi_\mu$. They are defined by

$$m = k_\mu'(\theta) = \int_{\mathbb{R}^n} xP(\theta, \mu)(dx), \quad M_F = k_\mu'(\Theta(\mu)), \quad \theta = \psi_\mu(m).$$

Note that since $k_\mu$ is strictly convex, then $k_\mu'$ is injective on $\Theta(\mu)$, the map $\psi_\mu$ from $M_F$ onto $\Theta(\mu)$ is well defined and $M_F$ is a connected open set.

If $C_F$ is the closed convex set generated by the support of $\mu$ clearly $M_F$ is contained in $C_F$. We say that $F$ or $\mu$ are steep if $M_F$ is equal to the interior of $C_F$. Most of the classical exponential families are steep, but not always (see for instance the Tweedie scale below for $p < 0$). In $\mathbb{R}$ the set $M_F$ is an interval, but in $\mathbb{R}^n$ there are non steep examples such that $M_F$ is non-convex (14, p. 35).

Finally, the last important object about the exponential family $F$ is its variance function $V_F$ defined on $M_F$ by

$$V_F(m) = \frac{1}{\psi_\mu'(m)} = k_\mu''(\psi_\mu(m)) = \int_{\mathbb{R}^n} (x - m) \otimes (x - m) P(\psi_\mu(m), \mu)(dx),$$

which characterizes $F$.

Now bilateral Laplace transforms are considered. In the particular case of dimension one, an older tradition associates the name of Laplace to integrals $\int_0^\infty e^{-sx} \mu(dx)$ which are conveniently defined for $s > 0$ in many circumstances. In the present paper we need to consider what the physicists call the bilateral Laplace transform

$$B_\mu(s) = \int_{\mathbb{R}^n} e^{-sx} \mu(dx).$$

Dealing with this slight change of notation $B_\mu(s) = L_\mu(-s)$ will much simplify the description of duality between two natural exponential families. In the sequel, we say 'Laplace transform' for short and by abus de langage instead of the longer term 'bilateral Laplace transform'.

Because we will deal with these bilateral Laplace transforms, we have to modify the description of the classical objects associated to $F = F(\mu)$ with

$$S(\mu) = -\Theta(\mu), \quad \ell_\mu(s) = k_\mu(-s), \quad m = -\ell_\mu'(s), \quad s = \psi_\mu(m) = -\psi_\mu(-m), \quad \ell_\mu''(s) = k_\mu''(-s), \quad V_F(m) = -(\psi_\mu'(m))^{-1}.$$ (4) (5)

Let us insist on the fact that $\ell_\mu$ is convex and thus $s \mapsto m = -\ell_\mu'(s)$ is decreasing when $n = 1$. For any $n$ the inverse function $m \mapsto s = \psi_\mu(m)$ from $M_F$ onto $S(\mu)$ is well defined. Thus we have several ways of coding a member of an exponential family:

$$P(\theta, \mu), \quad P(-s, \mu), \quad P(\psi_\mu(m), \mu), \quad P(-\psi_\mu(m), \mu).$$

Finally, in one dimension, to find a generating measure $\mu$ of the exponential family $F$ from the knowledge of $V_F$, we proceed as follows:

$$ds = -\psi_\mu'(m)dm = \frac{dm}{V_F(m)} \Rightarrow -s = \int \frac{dm}{V_F(m)}.$$

The choice of the integration constants will change $\mu$ and is arbitrary. In practical cases, we choose these constants in order to get the simplest form of $\mu$. Therefore this choice may depend on aesthetic considerations. Here is an important example in dimension one:

**Example 1** (the Tweedie scale). This term describes a set of exponential families with variance functions of the form $V_F(m) = \frac{m^\lambda}{\Gamma(1+\lambda)}$ on $M_F = (0, \infty)$ where $\lambda > 0$ and $p \in \mathbb{R} \setminus [0, 1]$ (see 11, 12, 21). We are going to describe their densities or Laplace transforms. The limiting cases $p = 0$ and $p = 1$ correspond respectively to the Gaussian exponential family with fixed variance $\lambda$ where $M_F = \mathbb{R}$ and to the Poisson family and they need a special treatment.

3
1. The stable subordinator case \( p > 2 \). For simplification we introduce \( \alpha = \frac{p-2}{p} \in (0, 1) \). We write

\[
ds = -\frac{\lambda^{p-1}dm}{m^p} \Rightarrow s = \frac{1}{p-1} \left( \frac{\lambda}{m} \right)^{p-1} = m = \frac{\lambda^{1-(1/p)}}{(p-1)^{1/(p-1)}} = -\ell'_\mu(s).
\]

Here \( S(\mu) = (0, \infty) \). We finally obtain \( \ell'_\mu(s) = -\lambda(\frac{p-1}{p})^p \lambda^p \). This family is generated by a stable law of parameter \( \alpha \) with Lévy measure concentrated on \((0, \infty)\) of Type 1 (see Section 4.3 for the definition of a Lévy measure of an infinitely divisible probability and its type).

2. The gamma case \( p = 2 \). Similarly we obtain \( S(\mu) = (0, \infty) \), \( \ell'_\mu(s) = -\lambda \) and \( B_\mu(x) = 1/s^\beta \). This family is generated by the measure \( \mu(dx) \) with density \( x^{\beta-1}/\Gamma(\lambda) \) and is the family of gamma distributions with shape parameter \( \lambda \).

3. The Poisson -gamma case \( 1 < p < 2 \). For simplification denote \( \omega = \frac{2-p}{p-1} \in (0, \infty) \). A computation analogous to the case \( p > 2 \) leads to

\[
\ell'_\mu(s) = \lambda \left( \frac{p-1}{p} \right)^\beta \frac{1}{2 - p} \frac{1}{s^\gamma}, \quad \ell''_\mu(s) = -\lambda \left( \frac{p-1}{p} \right)^{\beta-1} \frac{1}{2 - p} \frac{1}{s^{\gamma+1}}
\]

and to \( S(\mu) = (0, \infty) \). This implies that the corresponding exponential family is the set of laws of \( X_1 + \cdots + X_N(t) \) where \( X_1, \ldots, X_n, \ldots \) are iid gamma distributed with shape parameter \( \beta \) and where \( N(t) \) is an independent Poisson distribution with mean \( \lambda = \frac{\omega^{1-p} \gamma^{-2}}{2} \).

4. The non steep stable case \( p < 0 \). For simplicity denote \( q = -p > 0 \) and \( \gamma = \frac{4t^2}{q \tau_1^2} \in (1, 2) \). Here \( S(\mu) = (-\infty, 0) \) and we obtain

\[
\ell'_\mu(s) = \lambda \left( \frac{q+1}{q} \right)^\gamma \frac{1}{q+2} (-s)^\gamma.
\]

The members of this family are stable laws with parameter \( \gamma \) with Lévy measure concentrated on \((-\infty, 0)\). The support of such a stable law is \( \mathbb{R} \) but \( M_F = (0, \infty) \) : this is an example of a non steep family and of an infinitely divisible distribution of Type 2 (see Section 4 and (21)).

3. Duality

3.1. Duality between measures

Given \( \mu \in \mathcal{M}(\mathbb{R}^n) \) the function \( s \mapsto -\ell'_\mu(s) = m \) maps \( S(\mu) \) onto the domain of the means \( M_F \). Its inverse \( m \mapsto s = \varphi_\mu(m) \) exists and maps \( M_F \) onto \( S(\mu) \). Suppose that there exists \( \mu^* \in \mathcal{M}(\mathbb{R}^n) \) such that

\[
-\ell''_\mu(-\ell'_\mu(s)) = s.
\]

If (7) holds we say that \( \mu^* \) is a dual measure of \( \mu \). Note that we say ‘a dual measure’ since \( \mu^* \) is unique only up to a multiplicative constant. However the associated exponential family \( F(\mu^*) \) will not change if \( \mu^* \) is replaced by \( C\mu^* \). Observe also that if \( \mu^* \) exists and if \( \mu \) is steep then \( \mu^* \) is also a dual measure of \( \mu^* \). In general, if \( \mu^* \) is bounded it is natural to choose the multiplicative constant such that \( \mu^* \) is a probability.

**Proposition 1.** Let \( \mu \in \mathcal{M}(\mathbb{R}^n) \) and suppose that there exists a dual measure \( \mu^* \). Then \( \ell''_\mu(m) = (V_{F(\mu)}(m))^{-1} \). Furthermore, if \( \mu \) is steep \( \ell''_\mu(m) = (V_{F(\mu)}(m))^{-1} \).

**Proof:** Since \( s = \varphi_\mu(-\ell'_\mu(s)) \) we get by definition that

\[
-\varphi'_\mu(-\ell'_\mu(s)) = -(-\ell''_\mu(s))^{-1} = (V_{F(\mu)}(-\ell'_\mu(s)))^{-1}, \quad \ell'_\mu(m) = (V_{F(\mu)}(m))^{-1}.
\]

The second formula is obtained by symmetry.

In the particular case of dimension one, we will have numerous examples of dual measures given below. Let us give the simplest now, based on the Tweedie scale. Essentially, duality exchanges cases 1 and 3.
Proposition 2. Let $1 < p < 2$ and $q$ defined by $\frac{1}{p} + \frac{1}{q} = 1$. Let $\mu$, defined by $\mathbf{6}$, generating the exponential family with variance function $V_F(m) = \frac{m}{2}$. Then there exists a dual measure $\mu^*$ and it generates the exponential family with variance function $V_{F^\gamma}(s) = (\frac{2}{q})^{y}s^q$.

Proof: Using formula $\mathbf{6}$, we get $\ell_{\mu^*}'(s) = \lambda(p - 1)^{-q}s^{-q}$. Since $p - 1 = p/q$ Proposition $\mathbf{1}$ gives the result.

Remark 1. For having a more symmetric form we can choose $\lambda$ such that

$$V_{F^\gamma}(m) = \frac{R^p}{q^p}m^p, \quad V_{F^\gamma}(s) = \frac{R^{-q}}{p^{-q}}s^q.$$  

(8)

Dual measures do not always exist. Consider, for instance, Case 4 of the Tweedie scale above generated by a stable distribution $\mu$ with parameter $\gamma \in (1, 2)$ with Lévy measure concentrated on $(-\infty, 0)$ and domain of the means $M_{F^\gamma} = (0, \infty)$. Therefore if $\mu^*$ was existing we would have for positive constants $C, C_1, C_2$:

$$\ell_{\mu^*}(s) = C(-s)^\gamma, \quad \ell_{\mu^*}'(s) = -C\gamma(-s)^{\gamma-1} = -m,$$

$$\ell_{\mu^*}'(m) = C_1m^{\frac{\gamma}{\gamma-1}}, \quad \ell_{\mu^*}(m) = C_2m^{\frac{\gamma}{\gamma-1}}.$$  

This would imply that $\mu^*$ would be a stable law with parameter $\gamma/(\gamma - 1) > 2$, which is impossible.

Moreover, it is well known that a probability $P \in \mathcal{M}(\mathbb{R})$ is infinitely divisible if and only if there exists $\nu \in \mathcal{M}(\mathbb{R})$ such that for $s \in S(P)$ we have

$$\ell_{\nu}'(s) = \int e^{-sx}\nu(dx).$$  

(9)

Extension of the definition of infinite divisibility to measures $\mu \in \mathcal{M}(\mathbb{R})$ (even the unbounded ones) is easy and is done in Letac $\mathbf{[18]}$ and the same remark about $\ell_{\nu}'(s)$ holds. Then we have the following simple proposition:

Proposition 3. Let $\mu \in \mathcal{M}(\mathbb{R})$ and assume that a dual measure $\mu^*$ does exist. Then $\mu^*$ is infinitely divisible if and only if $m \mapsto \frac{1}{V_{F^\gamma}(m)}$ is the Laplace transform of some measure $\nu \in \mathcal{M}(\mathbb{R})$.

Proof: This an immediate consequence of Proposition $\mathbf{1}$ and of $\mathbf{20}$.

Example 2 (Bernoulli distribution on $\pm 1$). Proposition $\mathbf{5}$ gives a powerful tool to check quickly whether a dual measure exists. Consider, for instance, the Bernoulli measure on $\pm 1$ with mean $m \in (-1, 1)$:

$$\mu = \frac{1}{2}((1-m)\delta_1 + (1+m)\delta_1).$$

Therefore the variance of $X \sim \mu$ is $V_{F^\gamma}(\mu) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 1 - m^2$. To prove that $\mu^*$ exists we observe that

$$\frac{1}{1-m^2} = \frac{1}{2} \left( \frac{1}{1-m} + \frac{1}{1+m} \right) = \frac{1}{2} \int_{\mathbb{R}} e^{-m|y|}dy$$

and we apply Proposition $\mathbf{5}$ to the bilateral exponential law $\nu(dy) = \frac{1}{2} e^{-|y|}dy$. One can prove that $B_{\mu^*}(m) = (1 + m)^{1+m}(1-m)^{1-m}$ on $S(\mu^*) = (-1, 1)$. We will describe $\mu^*$ later on in Section 4, the second formula in $\mathbf{13}$ and Section 5.

3.2. The role of linear transformations and the Jørgensen set

Consider the image $\mu_d$ of $\mu$ by the isomorphism $s \mapsto a(s)$ of $\mathbb{R}^n$ into itself. Here $a$ is an invertible matrix of order $n$. If $\mu$ has a dual $\mu^*$ then $(\mu^*)^{-1}$ is a dual of $\mu_d$. Let us give the details of a tedious calculation:

$$B_{\mu_d}(s) = \int_{\mathbb{R}^n} e^{-a(s,a(s))} \mu_d(dx), \quad \ell_{\mu_d}(s) = \ell_{\mu}(a^T s), \quad m = -\ell_{\mu}''(s) = -a^T \ell_{\mu}''(a^T s),$$

$$(a^{-1})^T m = -\ell_{\mu}'(a^{-1} m), \quad a^T s = \ell_{\mu}'((a^{-1})^T m), \quad s = (a^{-1})^T \ell_{\mu}'(a^{-1})^T m, \quad (a^{-1})^T \ell_{\mu}'(a^{-1})^T m, \quad \ell_{\mu_d}'(m) = \ell_{\mu_d}'((a^{-1})^T m), \quad B_{\mu_d}(m) = B_{\mu}((a^{-1})^T m) = \int_{\mathbb{R}^n} e^{-a(s,a(s))} \mu'(dy) = \int_{\mathbb{R}^n} e^{-(a(s,a))} \mu'(dy) = B_{\mu_d}(m).$$
Recall also that the image of $F = F(\mu)$ by the isomorphism $a$ satisfies

$$V_{F(\mu)}(m) = a^T V_F(a^{-1}m)a.$$  

If $\mu \in \mathcal{M}(\mathbb{R}^n)$ then the Jørgensen set $\Lambda(\mu)$ is the set of $\lambda > 0$ such that there exists $\mu_\lambda \in \mathcal{M}(\mathbb{R}^n)$ such that $S(\mu) = S(\mu_\lambda)$ and $(B_\mu)^\lambda = B_{\mu_\lambda}$. Denote $F_\lambda = F(\mu_\lambda)$. Recall that the set $(F_\lambda)_{\lambda \in \Lambda(\mu)}$ is called by Jørgensen an exponential dispersion model. It is not correct to think that if $\mu^*$ is a dual measure of $\mu$ then $\Lambda(\mu) = \Lambda(\mu^*)$. We will see in Section 5 that a dual measure $\mu^*$ of the Bernoulli distribution $\mu$ exists and is infinitely divisible. Therefore

$$\Lambda(\mu^*) = (0, \infty) \neq \Lambda(\mu) = [1, 2, \ldots].$$

However if $\lambda$ is both in $\Lambda(\mu)$ and $\Lambda(\mu^*)$, then a dual measure of $\mu_\lambda$ satisfies

$$B_{\mu(\lambda)}(m) = B_{\mu^*}(\frac{m}{\lambda}).$$ \hspace{1cm} (10)

3.3. Duality and the change of generating measure of an exponential family

Let us clarify now what happens to $\mu^*$ and to $F(\mu^*)$ when we replace $\mu(dx)$ by $\mu_1(dx) = e^{(s_0 + s)}\mu(dx)$. This is important, since $F(\mu) = F(\mu_1)$.

**Proposition 4.** Let $\mu \in \mathcal{M}(\mathbb{R}^n)$ and suppose that there exists a dual measure $\mu^*$. Let $\mu_1(dx) = e^{(s_0 + s)}\mu(dx)$. Then $\mu_1 = \mu^* + \delta_{s_0}$ is a dual measure of $\mu_1$. In particular, the elements of $F(\mu_1)$ are the translation by $s_0$ of the elements of $F(\mu^*)$. Symmetrically if $\mu_2 = \mu * \delta_{m_0}$ is a translation of $\mu$ by $m_0$ then $\mu_2(dx) = e^{(m_0 + s)} \mu^*(dy)$ is a dual measure of $\mu_2$. In particular $\mu(\lambda) = F(\mu^*)$.

We skip the proof, which is an immediate application of the definitions. Let us consider the simple example of the gamma exponential family $F$ with shape parameter $\lambda$, which is Case 2 in the Tweedie scale example in Section 2. The domain of the means is $M_F = (0, \infty)$ and its variance function is $V_F(m) = m^2/\lambda$. Applying Proposition 4 we write

$$e^{\ell_{\mu'}_m} = \frac{\lambda}{m^2}, \quad e^{\ell_{\mu'}_m} = -\frac{\lambda}{m} + s_0, \quad \ell_{\mu'} = -\lambda \log m + s_0m + s_1, \quad B_{\mu'}(m) = e^{\ell_{\mu'} + s_1} = e^\gamma \int_0^\infty e^{-(y-s_0)m} \frac{y^{\lambda-1}}{\Gamma(\lambda)} dy.$$  

If $s_0 = 0$ then $F(\mu^*)$ is the same gamma family. If $s_0 \neq 0$ this is a translation of this last family. Suppose that we have started from a translated gamma family with $M_F = (m_0, \infty)$ and $V_F(m) = \lambda/(m-m_0)^2$. For this case we get the same families $F(\mu^*)$: they do not depend on the particular translation by $m_0$.

3.4. Duality for $T$–exponential families

The preceding discussion shows that while the dual of a measure is well defined up to a multiplicative constant, this is not the case for an exponential family $F$, since changing the generating measure $\mu$ of $F$ into some $\mu_1$ implies that $F(\mu) = F(\mu_1)$ by definition, but possibly changes $F(\mu^*)$ into a translate $F(\mu_1^*)$. For this reason we coin the definition of a $T$–exponential family:

**Definition 1.** Given an exponential family $F$ the associated $T$–exponential family $TF$ is the union of all the translations of $F$. If $F = F(\mu)$ then

$$TF = \{P(-s, \mu) * \delta_{m_0} : s \in S(\mu), \ m_0 \in \mathbb{R}^n\}.$$  

Such a $\mu \in \mathcal{M}(\mathbb{R}^n)$ is called a generating measure of $TF$.

Note that $TF(\mu) = TF(\mu_1)$ if and only if there exists $s \in S(\mu), b \in \mathbb{R}^n, m_0 \in \mathbb{R}^n$ such that $\mu_1 = \delta_{m_0} * e^{-(s_0 + b)} \mu(dx)$ or

$$\ell_{\mu_1}(s) = -(s, m_0) + b + \ell_{\mu}(s + s_0).$$

We are now in position to clearly define the dual $TF^*$ of a $T$–exponential family $TF$, thanks to the following proposition:
**Proposition 5.** If $\mu$ is a generating measure of the T-exponential family $TF$ in $\mathbb{R}^n$ and if $\mu$ has a dual measure $\mu^*$ then any generating measure $\mu_1$ has also a dual and

$$TF(\mu^*) = TF(\mu_1^*)$$

Under these circumstances we denote $TF^* = TF(\mu^*)$.

For instance we have seen that if $TF$ is the T–exponential family of all the gamma distributions with shape parameter $\lambda$ augmented with all its translations, then $TF = TF^*$ is self dual. We will find a similar phenomena with the Gaussian distributions with variance one, then with two unexpected examples: below in Section 6 with the dilogarithm exponential family and the sinh $m$ family, and, as a generalization of the one dimensional gamma case, with the Wishart distributions with fixed shape parameter.

### 3.5. Duality and large deviations.

Let $F = F(\mu)$ be a real T–exponential family and let $m_0 < m$ two points of $M_F$. Let $X_1, \ldots, X_n, \ldots$ be independent random variables with the same distribution in $F$ with mean $m_0$. The theorem of large deviations, due to Cramér [6], says that for

$$h(m_0, m) = \int_{m_0}^{m} \frac{m - t}{V_F(t)} dt$$

we have

$$\left( \Pr \left( \frac{1}{n}(X_1 + \cdots + X_n) > m \right) \right)^{1/n} \xrightarrow{n \rightarrow \infty} e^{-h(m_0, m)}. \quad (11)$$

In the next proposition, we link $h(m_0, m)$ with $\mu^*$ as the rest of the Taylor expansion of $m \mapsto \ell_{\mu^*}(m)$ around $m_0$. It shows that $m \mapsto \exp h(m_0, m)$ defines a member of $TF^*$ when the distribution $P$ of $X_0$ is in $TF$. One can consider such a proposition as a kind of probabilistic interpretation of duality in dimension one.

**Proposition 6.** Let $TF$ be a real T–exponential family such that its dual $TF^*$ exists. Let $X_1, \ldots, X_n, \ldots$ be independent random variables with the same distribution $P \in TF$ and mean $m_0$. For $m_0 < m$ define

$$h(m_0, m) = - \lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr \left( \frac{1}{n}(X_1 + \cdots + X_n) > m \right).$$

Suppose that the dual probability $P^*$ of $P$ exists. Let $s_0 = \ell_{P^*}(m_0), c = \ell_{P^*}(m_0) - m_0 \ell_{P^*}(m_0)$ and $P^*_{m_0} = e^{c} P^* \delta_{m_0}$. Then

$$B(P^*_{m_0})(m) = \int_{\mathbb{R}} e^{-my} P^*_{m_0}(dy) = e^{h(m_0, m)}. \quad (12)$$

**Proof:** From Proposition 5 we have $\ell_{\mu^*}(m) = 1/V_{F^{\mu^*}}(m)$. Therefore by integration by parts

$$h(m_0, m) = \int_{m_0}^{m} (m - t) \ell_{\mu^*}(t) dt = (m_0 - m) \ell_{\mu^*}(m) + \ell_{P^*}(m) - \ell_{P}(m) = \ell_{P^*}(m) - m s_0 + c.$$

\[\square\]

### 3.6. Warnings for existence and non existence of the dual measures

To conclude Section 3, devoted to generalities on duality, we consider some devices for proving that a function is not a Laplace transform for the case of dimension 1. Other ones will appear in Section 7 for the $\mathbb{R}^n$ case.

It will be shown that the Vinogradov-Paris distribution has no dual. Vinogradov and Paris [22] in their Theorem 6 consider the exponential family $F$ on $(0, \infty)$ with variance function $2m^2/(1 - m^2)$ defined on $M_F = (0, 1)$. Now we show that it has no dual. Indeed if a dual measure $\mu^*$ exists we have from Proposition 6

$$\ell_{\mu^*}'(m) = \frac{1 - m^2}{2m^2} \Rightarrow B_{\mu^*}(m) = e^{-m^2/2}/\sqrt{m}.$$
The function \( e^{-m^2/2} \) is a Fourier transform, and \( 1/\sqrt{m} \) is a Laplace transform. Who wins? By the theorem of maximal analyticity of Laplace transforms (see Letac and Mora \[15\] and Kawata \[11\]) \( B_{\varphi}(m) \) is not only defined on \( M_F = (0, 1) \) but also on \( (0, \infty) \). However the second derivative of \( \log B_{\varphi}(m) \) is negative on \( m > 1 \) and this proves that \( B_{\varphi} \) cannot be the Laplace transform of a positive measure. For proving that some \( TF^* \) does not exists, this trick can be utilized when \( F \) is not steep.

Let us apply Proposition 5 to the probability distributions

\[
P_1(dx) = \frac{1}{2} e^{-|m|}dx, \quad P_2(dx) = \frac{dx}{2 \cosh(\frac{1}{2} \pi x)}
\]

and to a distribution \( P_3 \) described in Letac \[16\] generating the exponential family with variance function \( (1 + m^2)^{3/2} \). The probabilities \( P_1 \) and \( P_2 \) are called the bilateral exponential and the hyperbolic distribution. They have means \( m_0 = 0 \), and respective Laplace transforms

\[
B_1(s) = \frac{1}{1 - s^2}, \quad B_2(s) = \frac{1}{\cos s},
\]

which generate two exponential families with respective variance functions

\[
V_1(m) = m^2 + 1 + \sqrt{m^2 + 1}, \quad V_2(m) = m^2 + 1.
\]

Performing the calculation of \( h(0, m) = \ell_{\varphi}(m) \) for \( P_1, P_2, P_3 \), we obtain

\[
h_1(m) = 2m^2 + 1 - \sqrt{m^2 + 1} e^{\arctan^{-1} m} - \frac{m}{2} e^{\arctan m}, \quad h_2(m) = \frac{1}{\sqrt{m^2 + 1}} e^{\arctan m}, \quad h_3(m) = e^{m^2 + 1}.
\]

For seeing this we write using for instance

\[
h_1(m) = 1 + \frac{3m^2}{4} + o(m^5), \quad h_2(m) = 1 + \frac{m^2}{2} + \frac{m^4}{24} + \frac{m^6}{80} + \frac{283m^8}{3020} + o(m^9), \quad h_3(m) = 1 + \frac{m^2}{2} + o(m^5).
\]

This proves that \( \int_{\mathbb{R}} x^4 P_1(dx) = \int_{\mathbb{R}} x^4 P_2(dx) = 0, \int_{\mathbb{R}} x^8 P_3(dx) = 0 < 0 \): none of them is possible. Thus \( P_1^*, P_2^* \) and \( P_3^* \) do not exist. My thanks go to Lev Klebanov for suggesting this method for proving the lack of duality of some measures.

4. The variance function \( e^m \), the Landau law \( \varphi \) and the dual of Poisson family

In this section, we restrict ourselves to the one dimensional case.

4.1. The Landau distribution \( \varphi \)

We will study \( TF^* \) when \( TF \) is the set of Poisson distributions augmented with its translations. In Section 5 we are going to study the dual pairs \( (TF, TF^*) \) issued of the case where \( V_T(m) \) is a quadratic or a cubic polynomial. For studying their dual we will use a particular probability \( \varphi \) on \( \mathbb{R} \) called the Landau law, which can be defined by its Laplace transform defined on \( S(\varphi) = (0, \infty) \) by \( B_{\varphi}(s) = e^{-s^2/2} \).

As we are going to see, \( \varphi \) exists and has support. Sometimes the term of Landau distribution is given to the convolution \( \varphi * \delta_{-1} \), because its Laplace transform has the more elegant form \( s^2 \). It is the law of \( X \sim 1 \) when \( X \sim \varphi \). A detailed study of the Landau distribution can be found, for instance, in Marucho et al. \[20\].

The law \( \varphi \) generates the exponential family with variance function \( e^m \). For seeing this we write using for instance

\[
\ell_\varphi(s) = s \log s - s, \quad \ell'_\varphi(s) = \log s = -m, \quad \ell''_\varphi(s) = \frac{1}{s} = e^m.
\]

An important feature of \( \varphi \) for our purposes is the existence of the following dual \( \varphi^* : \)

\[
s = -\ell'_\varphi(-\ell''_\varphi(s)) = \log \ell'_\varphi(s), \quad \ell'_\varphi(s) = e^{-s}, \quad \ell''_\varphi(s) = e^{-s} - 1, \quad B_\varphi(s) = e^{s-1}.
\]

Thus a dual of the Landau distribution is the Poisson distribution with mean 1.
4.2. Existence of $\phi$ and parents

**Proposition 7.** For $s > 0$ and $R > 1$ we have

$$e^{-sx} = \exp \int_{0}^{\infty} (e^{-sx} - 1 + sx e^{-x}) x^{-2} dx,$$

(15)

$$\frac{e^s}{(s+1)^{s+1}} = \exp \int_{0}^{\infty} (1 - e^{-sx})(1 - e^{-x}) x^{-2} dx,$$

(16)

$$\frac{s}{(s+1)^{s+1}} = R \int_{R}^{R+1} \exp - \int_{0}^{\infty} (1 - e^{-sx})(1 - e^{-(s+1)x}) x^{-2} dx,$$

(17)

$$\frac{1}{2} \frac{(s+2)^{s+2}}{(s+1)^{s+1}} = \frac{1}{2} \int_{0}^{\infty} (1 - e^{-sx})(x - 1 + e^{-x}) x^{-2} dx.$$

(18)

**Proof:** We prove first that for $a > 0$ we have

$$F(a) = \int_{0}^{\infty} (e^{-u} - 1 + au - u^2) du = -1 - \log a.$$

For proving the second statement in (15) we apply the first one while replacing $s$ by $1 + s$ and $1 - s$, thus obtaining two integrals. In the second one we change $x$ into $-x$ and thereafter summing the two integrals yield the second expression in (15) plus the extra term

$$2 + \int_{\mathbb{R}} (e^{-|x|} - 1 - |x|e^{-|x|}) \frac{dx}{x^2} = 0,$$

which is obtained by an integration by parts.

For proving the first relation in (16) we write patiently

$$\log \frac{s}{(s+1)^{s+1}} = s + \int_{0}^{\infty} (e^{-sx} - 1 + sx e^{-x}) x^{-2} dx - (s+1) - \int_{0}^{\infty} (e^{-sx} - 1 + sx e^{-x} + xe^{-x}) x^{-2} dx$$

$$= -1 + \int_{0}^{\infty} (e^{-sx} - 1 - e^{-x}) x^{-2} dx = -1 + \int_{0}^{\infty} e^{-xt} \left(1 - e^{-x} + \frac{e^{-x}}{1 - e^{-x}}\right)(1 - e^{-x}) x^{-2} dx$$

$$= -1 + \int_{0}^{\infty} e^{-xt} - 1)(1 - e^{-x}) x^{-2} dx + \int_{0}^{\infty} (1 - \frac{e^{-x}}{1 - e^{-x}})(1 - e^{-x}) x^{-2} dx = -\int_{0}^{\infty} (1 - e^{-sx})(1 - e^{-x}) x^{-1} dx.$$

The last line uses the fact observed above that $F(1) = -1$. For proving the second statement in (16) we use the Frullani integral

$$\log(1 + s) = \int_{0}^{\infty} (1 - e^{-sx}) e^{-x} x^{-1} dx$$

(19)

that we add to $\log \frac{s}{(s+1)^{s+1}}$ for obtaining the desired result.

The proof of (17) follows the same lines: we start from the log of (15) and replace $s$ by $s + 1$ and $s + \frac{1}{2}$. We watch the coefficient $f(x)$ of $e^{-sx}$ defined by (17) and we rearrange the remainder terms. The integral

$$\int_{0}^{\infty} f(x) dx = \frac{\log R}{R - 1}$$

is computed by an integration by parts followed by the application of (19).

For proving (18) we combine (19) and the first statement in (16) applied to $s + 1$ instead of $s$:

$$\frac{1}{2} \left(\frac{(s+2)^{s+2}}{(s+1)^{s+1}}\right) = \frac{1}{2} \int_{0}^{\infty} (1 - e^{-sx})(x - 1 + e^{-x}) x^{-2} dx + \int_{0}^{\infty} (1 - e^{-sx}) x^{-2} dx.$$

To conclude we have by an integration by parts and (19) that $(\frac{e^{-x}}{x})^2 dx = \log 4$. This ends the proof of (18).
4.3. Lévy measures.

Recall that we concentrate on the one dimensional case in this section. Extension to the \( \mathbb{R}^n \) case is possible, but we will have no use of it in this paper. Recall also that if a probability \( P \) is in \( \mathcal{M}(\mathbb{R}) \) then \( P \) is infinitely divisible if there exists a positive measure \( \nu(dx) \) on \( \mathbb{R} \setminus \{0\} \) such that

\[
\int_{\mathbb{R}\setminus\{0\}} \min(1, x^2) \nu(dx) < \infty,
\]

and there exist two numbers \( a \in \mathbb{R} \) and \( \sigma \geq 0 \) such that for all \( s \in S(P) \) we have

\[
\ell_P(s) = as + \frac{1}{2} \sigma^2 s^2 + \int_{\mathbb{R}\setminus\{0\}} (e^{-sx} - 1 + s \tau(x)) \nu(dx).
\]

(20)

Here \( \tau \) is a bounded function such that \( \tau(x)/x \xrightarrow{x \to 0} 1 \). Feller \( [7] \) chooses \( \tau(x) = \sin x \), the Russian school chooses \( \tau(x) = x/(1 + x^2). \) For getting simple formulas here we have chosen \( \tau(x) = x e^{-|s|} \). Changing \( \tau \) may change \( a \) but neither \( \sigma \) nor the measure \( \nu \), which is called the Lévy measure of \( P \). The number \( \sigma \) is called the Gaussian part. If \( \nu \) is bounded, we say that \( P \) or \( \nu \) is Type 0. If \( \nu \) is unbounded, but \( \int_{\mathbb{R}\setminus\{0\}} \min(1, |x|) \nu(dx) < \infty \), we say that \( P \) is Type 1. For Types 0 and 1, the representation of the Laplace transform \( B_P(s) \) does not need \( \tau \) and we can write

\[
\ell_P(s) = as + \frac{1}{2} \sigma^2 s^2 - \int_{\mathbb{R}\setminus\{0\}} (1 - e^{-sx}) \nu(dx).
\]

In other cases we say that \( P \) is of Type 2 and \( \tau \) is necessary. The convex hull of the support of such a probability of Type 2 is always \( \mathbb{R} \).

Example 3. In Section 2 we have mentioned in the Tweedie scale the unsymmetric stable distribution with parameter \( \gamma \in (1, 2) \) defined by its Laplace transform

\[
e^\gamma = \exp \left( as + \int_0^\infty (e^{-sx} - 1 + sxe^{-x}) x^{-\gamma-1} \frac{dx}{\gamma(\gamma-1)\Gamma(2-\gamma)} \right).
\]

(21)

This formula can be guessed by differentiating \( s^\gamma \) twice, and can be checked by two integrations by parts, yielding also the exact value of the constant \( a \). The Lévy measure is of Type 2 and equals

\[
\nu(dx) = 1_{(0,\infty)}(x) x^{-\gamma-1} \frac{dx}{\gamma(\gamma-1)\Gamma(2-\gamma)}.
\]

From Proposition \( [7] \) the first formula in \( [15] \), we see that the Landau distribution \( L \) exists, is infinitely divisible without Gaussian part and has \( \nu(dx) = 1_{(0,\infty)}(x) \frac{dx}{s^{\psi}} \) for Lévy measure. It is of Type 2.

From Proposition \( [7] \) the second formula in \( [15] \), we see that there exists an infinitely divisible distribution \( P \) such that \( B_P(s) = (1 + s)^{1+\gamma}(1 - s)^{-\gamma} \) with \( S(P) = \mathbb{R} \) and with Lévy measure

\[
\nu(dx) = e^{-|s|} x^{\gamma-1} \frac{dx}{\gamma(\gamma-1)\Gamma(2-\gamma)}.
\]

Its type is 2. We will see in Section 6 that \( P \) is the dual of the symmetric binomial distribution \( \frac{1}{2} (\delta_{-1} + \delta_1) \) as expected from \( [1] \) of the introduction.

From Proposition \( [7] \) the first formula in \( [16] \), we see that there exists an infinitely divisible distribution \( P \) such that \( B_P(s) = s^\psi/(s+1)^{\psi+1} \) with \( S(P) = (0, \infty) \) and with Lévy measure

\[
\nu(dx) = 1_{(0,\infty)}(x) \frac{1 - e^{-x}}{x} dx.
\]

Its type is 1. We will see later that \( P \) is the dual of the negative binomial distribution.
From Proposition 7 the second formula in (16), we see that there exists an infinitely divisible distribution $P$ such that $B_P(s) = s/(s + 1)^2$ with $S(P) = (0, \infty)$ and with Le\'vy measure

$$v(dx) = 1_{(0,\infty)}(x) \frac{1 - e^{-x}(1 + x)}{x^2} dx.$$ 

Its type is 0. We will see later that $P$ is the dual of an Abel distribution sometimes called generalized Poisson distribution, generating an exponential family on non-negative integers with cubic variance function $m(1 + m)^2$.

From Proposition 7 the first formula in (17) we see that there exists an infinitely divisible distribution $P$ such that $B_P(s) = \frac{\sum_{l=1}^{s}R_l}{(R + 1)\overline{R}}$ with $S(P) = (0, \infty)$ and with Le\'vy measure

$$v(dx) = 1_{(0,\infty)}(x) f(x) dx$$

where the positive function $f$ is defined by (17). Its type is 0, since the integral of $f$ is finite. We will see later that $P$ is the dual of a Tak\'acs distribution, generating an exponential family on non-negative integers with cubic variance function $m(1 + m)(1 + Rm)$.

From Proposition 7 the second formula in (17) we see that there exists an infinitely divisible distribution $P$ such that $B_P(s) = \frac{\sum_{l=1}^{s}R_l}{(R + 1)\overline{R}}$ with $S(P) = (0, \infty)$ and with Le\'vy measure

$$v(dx) = 1_{(0,\infty)}(x) \frac{x - 1 + e^{-x}}{x^2} dx.$$ 

Its type is 0. We will see later that $P$ is the dual of a Kendall-Ressel distribution, generating an exponential family with cubic variance function $m^2(1 + m)$.

5. A dictionary of dual pairs for quadratic and cubic families in $\mathbb{R}$

We split the section in two parts: duals of Morris families, duals of cubic families. The case of the Tweedie scale has been already ruled out in Section 3.1 and in particular, by Proposition 2. Most of the considered distributions are infinitely divisible, which is saying that the corresponding Jørgensen set is $(0, \infty)$. Formulas for passing from $\mu$ to $\mu_1$ and from $\mu_1$ to $(\mu_1)^*$ have been given above in Section 3.2. For this reason, except in the binomial case, we give simplified versions, ignoring the Jørgensen parameter $A$. For instance, looking at the dual of the negative binomial family with variance function $m + \frac{m^2}{\lambda}$, we deal only with $m + m^2$ in order to have more readable formulas for the duals.

The presentation of each pair $(TF, TF^*)$ is done by selecting a $\mu$ and a $\nu$ which may be only a translation of $\mu^*$. These measures are generating $TF$ and $TF^*$ respectively, namely $F(\mu) < TF$, $F(\nu) < TF^*$. We give one of the two variance functions $V_{TF}(\mu)$ and $V_{TF}(\nu)$ as well one of the two Laplace transforms $B_{TF}(\mu)$ and $B_{TF}(\nu)$. In most of the cubic cases, the densities on $\mathbb{R}$, or the weights in the discrete cases, are not available and either the Laplace transform or the variance function is computable explicitly: hence question marks may appear in the description of the dual.

Logic would have imposed to write $V_{TF}(\mu)$ and $B_{TF}(\mu)$ systematically. However we do not stick always to this rule. There are two reasons: variance functions of exponential families have become familiar objects during the last forty years from Morris [12], with many examples, and we are used to write $m$ not $s$ for the mean, exactly like we are reluctant to call $c$ the radius of a circle and $r$ its center. The second reason is after all that $TF^{**} = TF$. Which is the first, $TF$ or $TF^*$?

5.1. Duals of Morris families

1. The Poisson case

$$(m, \exp(e^{-s}); e^m, e^{-s} s).$$

The Poisson exponential family is generated by $\mu = \sum_{m=0}^{\infty} \frac{\delta_m}{m!}$ with Laplace transform $\exp(e^{-s})$. We have seen in Section 3.1 that $\ell_P(m) = \log m - m$ and therefore $\ell_P'(m) = \log m = -s$, $m = e^s$, $\ell_P''(s) = -e^{-s}$ and, ignoring the integration constant, $\ell_P''(s) = e^{-s}$. Therefore $\mu$ is a dual of the Landau distribution $\varphi$. 

11
2. **The Bernoulli case**

\[(m - m^2, 1 + e^{-s}; 4(\cosh \frac{m}{2})^2, s^2(1 - s)^{-1}).\]

The Bernoulli distribution is generated by \(\mu = \delta_0 + \delta_1\) leading to

\[\ell(\mu) = \log(1 + e^{-\mu}), \quad \ell'(\mu)(s) = -\frac{1}{1 + e^s} = -m,\]

where \(m\) is in the domain of the means \((0, 1)\) of the Bernoulli distribution. Therefore \(\ell_{\mu}(m) = -s = -\log(1 - m) + \log m\) and, ignoring the integration constant:

\[\ell_{\mu}(m) = (1 - m) \log(1 - m) + m \log m, \quad B_{\mu}(m) = m^m(1 - m)^{1-m}.\]

One can describe \(\mu^\ast\) by introducing the law \(\tilde{\varphi}\) of \(X\) when \(-X\) has the Landau distribution \(\varphi\), thus with Laplace transform \(e^m(-m)^m\) defined for \(m < 0\). Now we introduce the measure \(\varphi_1(dx) = e^{s1}\varphi(dx)\) with Laplace transform, defined for \(m < 1\),

\[B_{\varphi_1}(m) = \int e^{-ms}\varphi_1(dx) = e \int e^{-ms1}\varphi(dx) = e^{s1}(1 - m)^{1-m}.\]

Note, by assuming \(m = 0\) in \(B_{\varphi}(m), \varphi_1\) is a probability. Finally from the observation of their Laplace transforms we get a dual probability of \(\mu\) as the convolution of \(\varphi\) and \(\varphi_1\). We obtain

\[\mu^\ast = \varphi \ast \varphi_1, \quad B_{\varphi^\ast}(m) = e^{-m}m^me^m(1 - m)^{1-m} = m^m(1 - m)^{1-m},\]

since the product of two Laplace transforms is well defined on the intersection of their existence domain (here \((-\infty, 1)\) and \((0, \infty)\). For computing the variance function we use Proposition\[1\] and \(24\):

\[\frac{1}{V_{\mu}(m)} = \ell''(m) = \frac{d}{dm} \left| -\frac{1}{1 + e^m}\right| = \frac{4}{\cosh^2 \frac{m}{2}}.\]

In Proposition\[7\] and the second statement in \(22\), we have considered the case with the symmetric binomial distribution \(\mu = \frac{1}{4}(\delta_{-1} + \delta_{1})\) which is an affine transformation of the ordinary Bernoulli distribution. We have seen that its dual exists and has Laplace transform \((1 - s)^{1/(1 + s)}\) and is infinitely divisible with Lévy measure of Type 2. Since \(\ell_{\mu}(s) = \log \cosh s\) we get the variance function with Proposition\[1\]

\[V(m) = \frac{1}{\ell''(m)} = \cosh^2 m.\]

3. **The binomial case**

\[(m - m^2, N, \frac{1}{N}(1 + e^{-s}); 4(\cosh \frac{m}{2})^2, s^2(N - s)^{-1}).\]

For describing the dual of the binomial distribution from the study of the Bernoulli case, enough is to use \(10\).

4. **The negative binomial case**

\[(m + m^2, (1 - e^{-s})^{-1}; 4(\sinh \frac{m}{2})^2, s^2(1 + s)^{-1-s}).\]

The negative binomial is generated by \(\mu = \sum_{n=0}^\infty \delta_n\) leading to

\[\ell(\mu) = -\log(1 - e^{-s}), \quad \ell'(\mu)(s) = -\frac{1}{e^s - 1} = -m,\]

where \(m\) is in the domain of the means \((0, \infty)\). Therefore \(\ell_{\mu}(m) = -s = -\log(1 + m) + \log m\) and, ignoring the integration constant

\[\ell_{\mu}(m) = -(1 + m) \log(1 + m) + m \log m, \quad B_{\mu}(m) = m^m(1 + m)^{-1-m}.\]

Proposition\[7\] the first relation in \(18\), has shown the existence of the probability \(\mu^\ast\). It is an infinitely divisible distribution of Type 1. Since \(\ell_{\mu}(s) = -\frac{1}{e^s - 1}\), the variance function of \(F(\mu^\ast)\) is easily obtained:

\[V_{F_{\mu^\ast}}(m) = \frac{1}{\ell''(m)} = 4(\sinh \frac{m}{2})^2.\]
5. The gamma case

\[ m^2 \left( \frac{1}{A}, s^{-1}, \frac{m^2}{A}, s^{-1} \right). \]

This self dual has been detailed in Section 3 as a particular case of the Tweedie scale.

6. The Gaussian case

\[ (\sigma^2, \exp(\frac{1}{2} \sigma^2 s^2); \sigma^{-2}, \exp(\frac{1}{2} \sigma^{-2} s^2)). \]

This case is very simple since \( \ell_{N(0,\sigma^2)}(s) = \frac{1}{2} \sigma^2 s^2 \), \( \ell_{N(0,\sigma^{-2})}(s) = \sigma^{-2} s \) and therefore \( \ell_{N(0,\sigma^2)}(s) = \sigma^2 s \).

7. The hyperbolic case We have seen in Section 3.6 with equation (13) that these laws have no dual.

5.2. Duals of cubic families

1. The Abel case

\[ (m(1 + m)^2; ?, ?, s^2/(s + 1)^2). \]

2. The Takacs case

\[ (m(1 + m)(1 + Rm); ?, ?, \frac{s(s + 1)\sqrt{m}}{(Rs + 1)\sqrt{m}}). \]

3. The Kendall Ressel case

\[ (m^2(1 + m); ?, ?, \frac{(s + 1)}{s})^{s+1}. \]

The above given three cases are similar. The densities of these three cubic families are explicit (see Letac and Mora [15]) but not the Laplace transforms. Similarly the variance function of the dual is not computable (except if we accept a description by a sum of a series obtained by the Lagrange formula). However using Proposition [11] and the relation \( \ell''_{\mu}(m) = 1/V_{\mu}(m) \) we can catch the Laplace transform of the dual. Therefore we apply this principle to the three functions

\[
\begin{align*}
\frac{1}{m(1+m)^2} & \quad \frac{1}{m(1+m)(1 + Rm)} & \quad \frac{1}{m^2(1 + m)}.
\end{align*}
\]

Fortunately these relations can be integrated twice leading to the three Laplace transforms appearing in the second expression of (16), (17) and (18), respectively. For seeing that these three duals do exist, the hard work has been done in Proposition [7] where it is shown that these three duals are infinitely divisible, as commented on in Section 4.3. About [18], observe that the Laplace transform \( (\frac{1}{s})^{s+1} \) of an unbounded measure \( \mu \) is changed by the substitution \( s \mapsto s + 1 \) into the Laplace transform \( (\frac{1}{s+1})^{s+2} \), then converted into the Laplace transform \( (\frac{1}{s+1})^{s+2} \) of a probability of \( F(\mu) \):

1. The inverse Gaussian case

\[ (\frac{m^3}{A}, e^{-\frac{A^2}{2}}; 2^{3/2} \sqrt{m} 3^{1/2}, \exp \frac{A^2}{4s}). \]

This is an application of Proposition [11] with \( q = 3 \).

2. The strict arcsine case

\[ (m(1 + m^2); ?, ?, (\frac{s}{\sqrt{1 + s^2}})^{s} e^{\text{arctan}(s)}). \]

For computing the Laplace transform of \( \mu^* \) we use Proposition [11]

\[ \ell_{\mu^*}(s) = \frac{1}{s(1 + s^2)} = \frac{1}{1 + s^2} - \frac{s}{1 + s^2}, \quad \ell_{\mu^*}(s) = \log s - \frac{1}{2} \log(1 + s^2), \]

leading to \( B_{\mu}(s) = (\frac{1}{\sqrt{1 + s^2}})^{s} e^{\text{arctan}(s)} \). For seeing that the dual exists we note that

\[
\int_0^\infty e^{-sx}(1 - \cos x)dx = \frac{1}{s} - \frac{s}{1 + s^2} \ell_{\mu^*}(s) = \int_0^\infty (e^{-sx} - 1) \frac{1 - \cos x}{x^2}dx.
\]

Therefore \( \mu^* \) exists, is infinitely divisible with Lévy measure \( \frac{1 - \cos x}{x^2} \mathbf{1}_{(0,\infty)}(x)dx \) and is of Type 0.
3. The large arcsine family case \((m(1 + 2m \cos a + m^2), ?, ?, B_{\mu}(s))\) with \(0 < a < \pi/2\). The Laplace transform of the dual Proposition is obtained as follows:

\[
\ell_{\mu}(s) = \frac{1}{s(1 + 2s \cos a + s^2)} = \frac{1}{s} - \frac{1}{\sin^2 a} \left(\frac{s + 2 \cos a}{\sin a}\right)^2 + 1
\]

Observing that \(\ell_{\mu}(s)\) is the Laplace transform of a positive measure is not quite obvious. Introducing \(u = \frac{e^{-x \cos a}}{\sin a}\) leads to \(s = u \sin a - \cos a\) and to

\[
\frac{1}{\sin^2 a} \left(\frac{s + 2 \cos a}{\sin a}\right)^2 + 1 = \frac{1}{\sin^2 a} \left(\frac{u \sin a + \cos a}{u^2 + 1}\right) = \frac{1}{\sin^2 a} \left(\int_0^\infty e^{-\sin(a + t)} dt\right)
\]

\[
= \frac{1}{\sin^2 a} \left(\int_0^\infty e^{-x \cos a \sin(a + x \sin a)} dx\right)
\]

We now show that for \(x > 0\) and \(0 \leq a \leq \pi/2\) we have

\[
\frac{1}{\sin a} e^{-x \cos a \sin(a + x \sin a)} \leq 1
\]

or, equivalently, that \(f(x) = \sin x e^{x \cos a} - \sin(a + x \sin a) \geq 0\). For seeing this we use the two inequalities \(e^c \geq 1 + c\) and \(r \geq \sin r\) respectively applied to \(c = x \cos a\) and \(r = x \sin a\) and we get

\[
f(x) \geq \sin a(1 + x \cos a) - \sin a \cos r - \sin r \cos a = \sin a(1 - \cos r) + \cos a(r - \sin r) \geq 0.
\]

As a consequence, denoting

\[
g(x) = 1 - \frac{1}{\sin a} e^{-x \cos a \sin(a + x \sin a)} \sim x^2 / 2
\]

we can claim that \(\mu^*\) exists, is infinitely divisible with Lévy measure \(\frac{1}{\sqrt{2\pi}} 1_{(0,\infty)}(x) dx\) and is Type 0. One can note that the limit values \(a = \pi/2\) and \(a = 0\) yield the Abel and the strict arcsine cases. The computation of the function \(B_{\mu}(s)\) is a painful exercise of calculus. We get indeed

\[
\ell_{\mu}(s) = s \log s - \frac{1}{2}(s + \cos a) \log(s + \cos a) + \left(-\frac{1}{2} + (\frac{s}{2} - a) \cos a\right)s - \frac{1}{2}(\cotan a) \log(1 + 2s \cos a + s^2)
\]

\[
+ (\cotan a) \left(\frac{s + \cos a}{\sin a}\right) \arctan \left(\frac{s + \cos a}{\sin a}\right) + \frac{1}{2} \cos a \log \cos a + (\cotan a)^2 (\frac{s}{2} - a).
\]

6. The variance function \(e^m - 1\) and the dilogarithm law

In this section we still continue our study of duality in the one dimensional case. We introduce the dilogarithm law which generates the exponential family with variance function \(e^m - 1\). This gives the opportunity to introduce parent distributions \(\mu, \sigma, \sigma_r, \eta\) and \(\alpha\) with interesting properties, and exponential families with unexpected variance functions, like \(e^m + 1\) and sinh \(m\). We will see that \(\mu\) and \(\alpha\) are self dual.

6.1. The generating probability \(\mu\)

From the Bar Lev criteria described in Letac and Mora, \(V(m) = e^m - 1\) defined on \((0, \infty)\) is the variance function of an exponential family \(F\) concentrated on the non-negative integers. Let us compute a particular generating probability \(\mu = \sum_{n=0}^{\infty} \mu(n) \delta_n\) of \(F\). We use

\[
d s = -\frac{dm}{V(m)} = -\frac{dm}{e^m - 1} = -\frac{e^m dm}{1 - e^{-m}} = -d \log(1 - e^{-m}).
\]

Therefore \(S(\mu) = (0, \infty)\) and

\[
e^{-s} = 1 - e^{-m} \Rightarrow \ell_{\mu}(s) = -m = \log(1 - e^{-s}) \Rightarrow \ell_{\mu}(s) = -\sum_{n=1}^{\infty} \frac{e^{-mn}}{n}.
\]
Recall that
\[
\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.
\]
Continuing the calculation and choosing properly the integration constant such that \( k_\mu(0) = 0 \) we have
\[
\ell_\mu(s) = \frac{\pi^2}{6} + \sum_{n=1}^{\infty} \frac{e^{-ns}}{n^2} \Rightarrow B_\mu(s) = \sum_{n=0}^{\infty} \mu(n)e^{-ns} = \exp \left( \frac{-\pi^2}{6} + \sum_{n=1}^{\infty} \frac{e^{-ns}}{n^2} \right).
\]
As a consequence
\[
\sum_{n=0}^{\infty} \mu(n)z^n = \exp \left( \sum_{n=1}^{\infty} \frac{z^n - 1}{n^2} \right).
\] (26)
This generating function is linked to the classical special function \( L_2(z) \), called dilogarithm, defined on the unit disk \( \{ z \in \mathbb{C} : |z| \leq 1 \} \), for which a vast literature exists from Euler: let us quote Lewin [19], Zagier [24] and Kirilov [12]. The measure \( \mu \) is a probability as we can see by putting \( z = 1 \) in (26). It is an infinitely divisible distribution with Lévy measure \( \sum_{n=0}^{\infty} \frac{\delta_n}{\pi}. \) For this reason we call \( \mu \) the dilogarithm law in the sequel. Here is a surprising property that \( \mu \) shares with the normal and the gamma laws, namely the self duality.

**Proposition 8.** The dilogarithm distribution \( \mu \) is dual of itself.

**Proof:** The result is immediate from (25) which implies the symmetric relation \( e^{-s} + e^{-m} = 1. \)

Consider \( Y \) and \( Y' \) which are independent with the same distribution \( \mu \). We denote by \( \sigma \) the distribution of \( Y - Y' \) in \( \mathbb{Z} \). Consider also an element \( P(-s, \mu) \) of the exponential family \( P(\mu) \). Since \( S(\mu) = (0, \infty) \) we rather write \( r = e^{-s} \in (0, 1) \) and we consider the probability \( \mu_r = P(-s, \mu) \) which satisfies
\[
\sum_{n=0}^{\infty} \mu_r(n)z^n = e^{L_2(rz) - L_2(r)}.
\] (27)
Similarly, we introduce \( Y_r \) and \( Y'_r \) independent with the same distribution \( \mu_r \) and denote by \( \sigma_r \) the distribution of \( Y_r - Y'_r \) in \( \mathbb{Z} \). In Sections 6.2 and 6.3 we study the four probabilities \( \mu, \sigma, \mu_r, \sigma_r. \)

### 6.2. Some properties of \( \mu \) and \( \sigma \)

For having a probabilistic interpretation of the probability \( \mu \) defined by (26), consider the probability \( \nu \) on positive integers defined by
\[
\nu = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{\delta_n}{n^2}.
\]
Suppose that \( X_1, \ldots, X_N \) are iid with distribution \( \nu \) and consider an independent Poisson random variable \( N \) with mean \( \lambda = \pi^2/6 \). Then the distribution of \( X_1 + \cdots + X_N \) is \( \mu \) by a classical calculation using conditioning by \( N \). Of course, this shows that \( \mu \) is infinitely divisible of type 0.

**Proposition 9.** If \( Y \) and \( Y' \) are independent with the same distribution \( \mu \) then the characteristic function of the random variable \( Y - Y' \sim \sigma \) on \( \mathbb{Z} \) is the function with period \( 2\pi \) given for \( 0 \leq t \leq 2\pi \) by \( \mathbb{E}(e^{it(Y-Y')}) = e^{-\frac{1}{2}t(2\pi-t)}. \)

**Proof:** Since \( \mathbb{E}(e^{itY}) = \exp \left( \sum_{n=1}^{\infty} \frac{e^{int} - 1}{n} \right) \) we have
\[
\mathbb{E}(e^{it(Y-Y')}) = \exp \left( 2 \sum_{n=1}^{\infty} \frac{\cos nt - 1}{n} \right) = \exp \left( -\frac{1}{2}t(2\pi - t) \right),
\]
and due to elementary calculation of Fourier series we have, for \( 0 \leq t \leq 2\pi \), \( \sum_{n=1}^{\infty} \frac{\cos nt}{n} = \frac{\pi^2}{6} - \frac{1}{2}t(2\pi - t). \)
Remark 2. 1) Since $S(\mu) = (0, \infty)$, then $\sigma$ has no Laplace transform and cannot generate an exponential family. 2) The calculation of $\Pr(Y = Y')$ is not elementary. More specifically by the change of variable $s = t - \pi$ and Mathematica:

\[
\Pr(Y = Y') = \frac{1}{2\pi} \int_0^{2\pi} e^{-i\theta(2\pi - \pi)} d\theta = \frac{1}{\pi} \int_0^{\pi} e^{i(2\pi - \pi)^2} d\theta = 0.11751.
\]

By a similar calculation for $n \in \mathbb{Z}$

\[
\Pr(Y - Y' = n) = \frac{1}{2\pi} \int_0^{2\pi} e^{-i\theta(2\pi + n)} d\theta = \frac{(-1)^n}{\pi} \int_0^{2\pi} e^{i(2\pi + n)^2} \cos(ns) ds.
\]

The fact that $E(e^{r(Y - Y')}) = e^{\frac{r^2}{2}}$ in $[0, 2\pi]$ calls for a link with the Gaussian distribution, as it is shown in the next proposition.

Proposition 10. If $Z \sim N(0, 1)$ and $Y, Y' \sim \mu$ are independent, then the logarithm of the characteristic function of the continuous random variable $W = Z + Y - Y'$ is made of piecewise affine functions and is concave. More specifically, if $t - 2\pi k \in [0, 2\pi)$ with $k \in \mathbb{Z}$ then

\[
\log E(e^{iW}) = 2k(k + 1)\pi^2 - (2k + 1)\pi t
\]

and the density $f(x)$ of $W$ is, using the notation $e_k(x) = e^{-2\pi^2 x^2 - 2ixs}$, equals

\[
f(x) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e_k(x) - e_{k+1}(x).
\]

(28)

Proof: If $t - 2\pi k \in [0, 2\pi)$ with $k \in \mathbb{Z}$ we have

\[
\log E(e^{iW}) = -\frac{1}{2}t^2 + (t - 2\pi k)^2 - \pi(t - 2\pi k) = 2k(k + 1)\pi^2 - (2k + 1)\pi t.
\]

To compute $f(x)$ we use the inverse Fourier transform leads easily to (28):

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} E(e^{itW}) dt = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{2\pi k}^{2\pi (k+1)} e^{-itx + 2k(k+1)\pi^2} dt.
\]

6.3. Some properties of $\mu$ and $\sigma$

Proposition 11. Let $0 < \rho < 1$. If $Y_r$ and $Y'_r$ are independent with the same distribution $\mu$, then the Laplace transform of the law $\sigma_r$ of $Y_r - Y'_r$ is defined for $s \in S(\sigma_r) = (\log r, \log(1/r))$ by

\[
E(e^{-s(Y_r - Y'_r)}) = \exp \left(-\int_0^1 \log(1 - 2x \cosh \theta + x^2)\frac{dx}{x} - 2\int_0^1 \log(1 - x)\frac{dx}{x} \right).
\]

The distribution $\sigma_r$ generates an exponential family $F(\sigma_r)$, where $\mathbb{R}$ is the domain of the means and the variance function is, with $a = \sinh \frac{\pi}{2}$.

\[
V_{F(\sigma_r)}(m) = \frac{2}{1 - r^2 \sqrt{r^2 + a^2}} \left( \sqrt{r^2 + a^2} + \sqrt{1 + a^2} \right).
\]

Proof: From (27) $B_{\sigma_r}(s)$ exists if and only if $re^{-s} < 1$ and $re^{-s} < 1$, in other terms if $|s| < -\log r$. Since $L_{2s}(r) = -\int_0^1 \log(1 - x)\frac{dx}{x}$ we get

\[
L_{2}(re^{-s}) = -\int_0^{re^{-s}} \log(1 - x)\frac{dx}{x} = -\int_0^{\infty} \log(1 - xe^{-s})\frac{dx}{x}
\]

and $B_{\sigma_r}(s)$ is easily deduced from this expression.
For computing the variance function, we observe that
\[ m = k'_r(s) = \log(1 - re^{-s}) - \log(1 - re^s). \]  
(29)

Equality (29) shows that \( \mathbb{R} \) is the domain of the means by letting \( s \to \pm \log r \). Also, (29) can be rewritten,
\[ a = \sinh \frac{m}{2} = r \sinh \left( -s + \frac{m}{2} \right) \]  
(30)

With the classical notations \( \theta = \psi(m) \) and \( V_{F(r)}(m) = 1/\psi'(m) \) we can write, by differentiating (30) with respect to \( m \),
\[ \frac{1}{2} \sqrt{1 + a^2} = \frac{1}{2} \cosh \frac{m}{2} = r \cosh \left( \psi(m) + \frac{m}{2} \right) \left( \psi'(m) + \frac{1}{2} \right) = r \sqrt{1 + \frac{a^2}{r^2}} \left( \frac{1}{V_{F(r)}(m)} + \frac{1}{r} \right). \]  
This transform leads to the desired formula:
\[ V_{F(r)}(m) = \frac{2 \sqrt{r^2 + a^2}}{\sqrt{1 + a^2} - \sqrt{r^2 + a^2}} = \frac{2}{1 - r^2} \sqrt{r^2 + a^2} \left( \sqrt{1 + a^2} + \sqrt{r^2 + a^2} \right). \]

The calculation of the dual of \( \sigma_r \) involves elliptic integrals and its existence is unproven.

6.4. The variance function \( e^{m} + 1 \)

In this section we study the exponential family \( F(\eta) \) on \( \mathbb{R} \) with variance function \( e^{m} + 1 \) and with the most surprising property that probability \( \eta \) is equal to \( N(0, 1) + \mu \) where \( \mu \) is the dilogarithm probability defined by (26).

Since the function \( e^{m} + 1 \) is real analytic and positive on \( \mathbb{R} \), then, if it is a variance function, the domain of the means of the corresponding natural exponential family will be \( \mathbb{R} \). Now, to decide whether it is a variance function or not, we have to compute a potential Laplace transform in the usual way:
\[ ds = -\frac{dm}{e^{m} + 1} = \frac{e^{-m} dm}{1 + e^{-m}} = -\frac{du}{1 + u} = -d \log(1 + u) = -d \log(1 + e^{-m}). \]

Since \( m > 0 \) we can take \( S(\eta) = (0, \infty) \) and \( s = \log(1 + e^{-m}) \). This leads to
\[ \ell'_\eta(s) = -m = \log(e^s - 1) = s + \log(1 - e^{-s}) = s - \sum_{n=1}^{\infty} \frac{e^{-ns}}{n}. \]

By choosing properly integration constant such that \( \eta \) is a probability, we get
\[ \ell_\eta(s) = \frac{1}{2} s^2 - \frac{1}{8} + \sum_{n=1}^{\infty} \frac{e^{-ns}}{n} = \ell_{N(0,1)}(s) + \ell_\mu(s). \]

Showing the unexpected result that \( \eta = N(0, 1) + \mu \) proves also that \( \eta \) exists and is a positive measure, and that \( e^{m} + 1 \) is the variance function of an exponential family. Proposition(11) given above can be reformulated in terms of \( \mu \) and \( \eta \). If \( X \sim \eta \) and \( Y \sim \mu \) are independent, then \( W \sim X - Y \). One can remark that since \( e^{m} \) and \( e^{m} + 1 \) are variance functions, therefore using affinities and the Jørgensen set would enable us to describe all exponential families with variance functions \( Ae^{m/1} + B \). We leave to the reader to prove that \( \eta \) has no dual.

6.5. The variance function \( \sinh(m) \)

In this section we study the exponential family \( F(\alpha) \) on \( \mathbb{N} \) with variance function \( \sinh(m) \) and we describe the link between the probability \( \alpha = (\alpha(n))_{n=0}^{\infty} \) and the dilogarithm function \( L_2(z) \). The existence of \( \alpha \) is granted by the Bar-Lev theorem since the Taylor expansion of \( \sinh(m) \) has only nonnegative coefficients (see [15], Corollary 3.3). The fact that \( \alpha \) is concentrated on \( \mathbb{N} \) is granted by the fact that \( \sinh(m) \) \( m \to 0 \) (see [15], Proposition 4.4). For computing \( \alpha \) we proceed in the usual way
\[ ds = -\frac{dm}{\sinh(m)} = -\frac{2e^{m} dm}{e^{2m} - 1} = -\frac{2du}{u^2 - 1} = -\left( \frac{1}{u - 1} - \frac{1}{u + 1} \right) du = -d \log \frac{u - 1}{u + 1} = -d \log \frac{e^m - 1}{e^m + 1}. \]

17
Since \( m > 0 \) we can take \( S(\alpha) = (0, \infty) \) and \( s = -\log \frac{e^{2\alpha}}{e^{2\alpha} - 1} \). This leads to
\[
\ell''(s) = -m = -\log \frac{1 + e^{-s}}{1 - e^{-s}} = -2 \sum_{n=1}^{\infty} \frac{e^{-2n-1} s^n}{2n - 1}.
\] (31)

Writing \( z = e^{-s} \in (0, 1) \) we get
\[
\ell''(s) = C + 2 \sum_{n=1}^{\infty} \frac{z^{2n-1}}{(2n - 1)^2} = C + 2 \text{Li}_2(z) - \frac{1}{2} \frac{\text{Li}_2(z^2)}{1 - x^2}.
\]

Since \( \text{Li}_2(1) = \pi^2/6 \), in order to have \( \alpha \) of mass 1, we take \( C = -\pi^2/4 \) and we finally obtain
\[
\sum_{n=1}^{\infty} \alpha(n)z^n = e^{-\frac{x^2}{2} + \frac{1}{2} \text{Li}_2(z^2)}.
\] (32)

Of course, if
\[
\beta(dx) = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{(2n - 1)^2} \delta_{2n-1},
\]
if \( N \) is Poisson distributed with mean \( \pi^2/4 \), and if \( X_1, \ldots, X_n, \ldots \) are iid with distribution \( \beta \) and independent of \( N \), we have \( X_1 + \cdots + X_N \sim \alpha \).

**Proposition 12.** The probability \( \alpha \) is dual of itself.

**Proof:** From (31) we obtain \( e^{-m} + e^{-s} + e^{s-m} = 1 \). The symmetry between \( m \) and \( s \) implies the result.

The last proposition links \( \alpha \) and \( \mu \). The proof follows from (32).

**Proposition 13.** If \( Y \sim \mu \) and \( X \sim \alpha 2 \) are independent, then \( X + 2Y \sim \mu^4 \).

### 7. Examples of duality in \( \mathbb{R}^n \)

The classification of exponential families in \( \mathbb{R}^n \) has been done in the literature with the same guidelines as for \( \mathbb{R} \): from the simplest variance functions like the Tweedie scale, or like the quadratic ones of Morris [21] to more complicated ones like the cubic families (see [15]) or the Babel class (see [16]). For \( \mathbb{R}^n \), several choices of classification through a definition of simple variances have been carried out:

1. The first one has been to extend the gamma family with shape parameter \( p > 0 \), whose variance function is \( m^2/p \), to homogeneous quadratic variance functions in \( \mathbb{R}^n \). More specifically one had to find which variance matrices \( V_{\mu}(m) \) of order \( n \) are made of homogeneous quadratic polynomials with respect to \( m = (m_1, \ldots, m_n) \). The answer has been given by Casalis [3]: out of trivial cases, the only exponential families with homogeneous quadratic variance functions are the Wishart ones, on symmetric, Hermitian, quaternionic Hermitian matrices, and analogous objects on the Lorentz cones and on the exceptional Albert algebra.

2. The second choice has been to consider the so called simple quadratic families in \( \mathbb{R}^n \). While a quadratic family has variance function of the form \( V_{\mu}(m) = A(m) + B(m) + C \), where \( m \mapsto A(m) \) is quadratic homogeneous, \( m \mapsto B(m) \) is linear and \( C \) is a constant symmetric matrix of order \( n \), such a family is said to be simple if \( A(m) \) has rank one, and therefore of the form \( \frac{1}{2} m \otimes m \); recall that if \( E \) is a Euclidean space and \( a, b \) are in \( E \) then \( a \otimes b \) is the endomorphism of \( E \) defined by
\[
x \mapsto (a \otimes b)(x) = a(b, x).
\]

Here again their classification has been done by Casalis [4], while the classification of general quadratic families is an open problem.

3. The third block is the simple cubic exponential families, obtained from the simple quadratic ones by a Möbius transformation of their variance function. They have been classified by Hassairi [3].

4. The last choice is the so called diagonal families in \( \mathbb{R}^n \) defined by the fact that the diagonal part of \( V_{\mu}(m) \) has the form \( (V_{\mu}(m_1), \ldots, V_{\mu}(m_n)) \). They have been classified in a paper with six coauthors: see Bar-Lev et al. [2].
7.1. The Wishart families are self dual

In the Euclidean space $V$ of real symmetric matrices of order $n$ with scalar product $\langle x, y \rangle = \text{trace}(xy)$ consider the cone $V_+ \subset V$ of positive definite matrices and the cone $\overline{V_+} \subset V$ of semipositive definite matrices. If

$$p \in \Lambda = \left\{ \frac{1}{2}, \frac{1}{2}, \ldots, \frac{n-1}{2} \right\} \cup \left( \frac{n-1}{2}, \infty \right),$$

the family of the Wishart distributions of shape parameter $p$ is generated by the unbounded measure $\mu_p$ on $\overline{V_+}$ which can be defined on $s \in V_+$ by its Laplace transform

$$\int_{V_+} e^{-\langle s, x \rangle} \mu_p(dx) = \left( \text{det} s \right)^p.$$

As a consequence $\ell_{\mu_p}(s) = -p \text{det} s$, $\ell'_{\mu_p}(s) = -ps^{-1}$ since the gradient of $s \mapsto \log \text{det} s$ is $s^{-1}$. Clearly $-\ell_{\mu_p}(-\ell'_{\mu_p}(s)) = p(p^{-1}s)^{-1} = s$ and this shows that $\mu_p$ is a dual of itself.

We are not going to give details for the Wishart distributions defined on the other symmetric cones: after introducing the necessary definitions, the simple calculation above remains the same. One can consult Casalis and Letac [5], for instance.

7.2. The dual of the multinomial distribution

Let $(e_1, e_2, \ldots, e_n)$ be the canonical basis of $\mathbb{R}^n$ and let us call Bernoulli distribution any law concentrated on the $n+1$ points $(0, e_1, e_2, \ldots, e_n)$, namely of the form

$$p_0\delta_0 + p_1\delta_{e_1} + p_2\delta_{e_2} + \cdots + p_n\delta_{e_n},$$

where $p_i > 0$ for $i \in [0, \ldots, n]$ and $p_0 + p_1 + \cdots + p_n = 1$. The set of all these Bernoulli distributions is an exponential family generated by the measure $\mu = \delta_0 + \delta_{e_1} + \cdots + \delta_{e_n}$ whose Laplace transform is

$$\Delta = B_p(s_1, \ldots, s_n) = 1 + e^{-s_1} + \cdots + e^{-s_n}.$$

Therefore

$$\ell_{\mu}(s) = \log \Delta, \quad m_1 = -\frac{\partial}{\partial s_1} \ell_{\mu}(s) = \frac{e^{-s_1}}{\Delta}, \ldots, m_n = -\frac{\partial}{\partial s_n} \ell_{\mu}(s) = \frac{e^{-s_n}}{\Delta}.$$

For deciding whether $\mu^*$ exists or not we compute $\ell_{\mu^*}(m)$ such that $-\ell_{\mu^*}(-\ell'_{\mu^*}(s)) = s$. Thus we have to compute $(s_1, \ldots, s_n)$ with respect to $(m_1, \ldots, m_n)$ as follows:

$$-m_1\Delta = e^{-s_1}, \ldots, -m_n\Delta = e^{-s_n}, \quad 1 + (m_1 + \cdots + m_n)\Delta = \Delta, \quad \Delta = \frac{1}{1 - m_1 - \cdots - m_n}.$$

Therefore

$$s_1 = -\log m_1 + \log(1 - m_1 - \cdots - m_n), \ldots, s_n = -\log m_n + \log(1 - m_1 - \cdots - m_n).$$

As a consequence, if $\mu^*$ does exist, it must satisfy

$$\ell'_{\mu^*}(m) = (\log m_1 - \log(1 - m_1 - \cdots - m_n), \ldots, \log m_n - \log(1 - m_1 - \cdots - m_n)).$$

It is easy to see that up to an additive constant we have

$$\ell_{\mu^*}(m) = (m_1 \log m_1 + \cdots + m_n \log m_n + (1 - m_1 - \cdots - m_n) \log(1 - m_1 - \cdots - m_n),$$

$$B_{\mu^*}(m) = m_1^{m_1} \cdots m_n^{m_n} (1 - m_1 - \cdots - m_n)^{1 - m_1 - \cdots - m_n}.$$

Recall that the Landau distribution $\varphi$ is defined by $\int_{\mathbb{R}^n} e^{-\langle x, x \rangle} \varphi(dx) = e^{-\frac{1}{2}x^2}$. In order to describe the above $\mu^*$ let us coin a simple lemma:
Lemma 1. Let \( f(s) = f(s_1, \ldots, s_n) = a_1s_1 + \cdots + a_ns_n + b = \langle a, s \rangle + b \) be a non-constant affine form of \( \mathbb{R}^n \). Then \( e^{-f(s)}f(s) \), defined on the half plane \( \mathbb{H} = \{ s; f(s) > 0 \} \), is the Laplace transform of a positive measure \( \mu(dx) \) on \( \mathbb{R}^n \), concentrated on the line \( \Re a \), defined as the image of the measure \( \varphi(dt) = e^{-\Re a} \varphi(dt) \) by the map \( t \mapsto x = at \). This measure is bounded if \( b \geq 0 \), of mass \( e^b b^n \) when \( b > 0 \), and mass 1 if \( b = 0 \).

Proof: By definition we have

\[
\int_{\mathbb{R}^n} e^{-\langle s, \lambda \rangle} \mu(dx) = \int_{\mathbb{R}} e^{-\langle s, \lambda \rangle} \varphi(dt) = \int_{\mathbb{R}} e^{-f(s)} \varphi(dt) = \int_{\mathbb{R}} e^{-f(s)} \varphi(dt).
\]

The remainder is plain.

With this lemma we can describe \( \mu^*(dx) \), which satisfies

\[
\int_{\mathbb{R}^n} e^{-\langle s, \lambda \rangle} \mu^*(dx) = s_1^{\prime} \cdots s_n^{\prime} (1 - s_1 - \cdots - s_n)^{1-s_1-\cdots-s_n}
\]
on the tetrahedron \( \{ s; s_1, \ldots, s_n, 1 - s_1 - \cdots - s_n > 0 \} \) which is the domain of the means of the exponential family of the Bernoulli distribution. From the lemma, applied to the \( n+1 \) affine forms \( f_0(s) = -s_1 - \cdots - s_n + 1, f_1(s) = -s_1, \ldots, f_n(s) = s_n \), the measure \( \mu^* \) is the convolution of \( n+1 \) probabilities respectively concentrated on the lines \( \Re e_1, \ldots, \Re e_n \) and \( \Re(-e_1 - \cdots - e_n) \). For \( f_0 \) the measure described in the lemma is bounded and has mass \( e^{-1} \) and has to be normalized to become a probability:

\[
e^{-f_0} e^{-f_1} \cdots e^{-f_n} = (1 - s_1 - \cdots - s_n)^{1-s_1-\cdots-s_n} s_1^{\prime} \cdots s_n^{\prime}.
\]

It is interesting to compute the variance function of \( F(\mu^*) \):

**Proposition 14.** Let \( \Delta = 1 + e^{-s_1} + \cdots + e^{-s_n} \), \( \mathbf{D} = \text{diag}(e^{-s_1}, \ldots, e^{-s_n}) \) and \( J_n \) the \((n, n)\) matrix with all entries equal to one. Let \( \mu^* \) be a dual of \( \mu = \delta_0 + \delta_{e_1} + \cdots + \delta_{e_n} \). Then the variance function of \( F(\mu^*) \) is

\[
\text{Var}_F(\mu^*)(s) = \Delta(D^{-1} + J_n) \left[ e^{s_1} + \cdots + e^{s_n} \right] \left[ \begin{array}{c} e^{s_1} + 1 \\ \vdots \\ e^{s_n} + 1 \end{array} \right].
\]

Proof: Let \( v = v(s) = (e^{-s_1}, \ldots, e^{-s_n})^T \). Note that \( \Delta' = v \) and that \( v' = -D \). With this notation we have, since \( \ell_\mu = \log \Delta \),

\[
\ell'_\mu(s) = -\Delta v, \quad \ell''_\mu(s) = \frac{1}{\Delta} D - \frac{1}{\Delta^2} v \otimes v, \quad V_{F(\mu^*)}(s) = [\ell''_\mu(s)]^{-1}.
\]

For computing the inverse matrix we use the following identity: if \( a \in \mathbb{R}^n \) such that \( \|a\|^2 = a_1^2 + \cdots + a_n^2 \) is different from 1, then

\[
(I_n - a \otimes a)^{-1} = I_n + \frac{a \otimes a}{1 - \|a\|^2}.
\]

We are going to apply this expression to

\[
a = \Delta^{-1/2} D^{-1/2} v = \Delta^{-1/2}(e^{-s_1/2}, \ldots, e^{-s_n/2})^T,
\]

which satisfies

\[
\|a\|^2 = \frac{\Delta - 1}{\Delta}, \quad 1 - \|a\|^2 = \frac{1}{\Delta}, \quad \frac{a \otimes a}{1 - \|a\|^2} = (D^{-1/2} v) \otimes (D^{-1/2} v),
\]

\[
V_{F(\mu^*)}(s) = \left( \frac{1}{\Delta} D - \frac{1}{\Delta^2} v \otimes v \right)^{-1} = \Delta D^{-1/2} (I_n - a \otimes a)^{-1} D^{-1/2} = \Delta D^{-1/2} (I_n + \frac{a \otimes a}{1 - \|a\|^2}) D^{-1/2} = \Delta D^{-1} + \Delta(D^{-1/2} a) \otimes (D^{-1/2} a) = \Delta(D^{-1} + J_n).
\]

\[\square\]

**Remark 3.** Note that when \( n = 1 \) the variance is \((1 + e^s)(1 + e^s) = 4 \cos^2 \frac{s}{2} \) as seen in (23) .
7.3. The other simple quadratic families.

Casalis [14] splits the set of simple quadratic families in $\mathbb{R}^n$ in $2n + 4$ types, extending the famous Morris classification for $n = 1$ in six families. Among them $n + 1$ are trivial from our point of view, with variance function of the form $B(m) + C$ where $m \mapsto B(m)$ is linear. In fact, with a proper choice of a basis of $\mathbb{R}^n$ they are product of $k$ Poisson families and $n - k$ one dimensional Gaussian families with $k \in \{0, \ldots, n\}$ (see also [14]). Since trivially a product of duals is dual of the product, there is nothing to add. The remaining $n + 3$ families are made with two exceptional ones, the multinomial and the hyperbolic, and a set of $n + 1$ types denoted $(\text{NM} - ga)_k$ families with $k \in \{0, \ldots, n\}$ that we will describe later. In the sequel we are going to prove that neither the hyperbolic type nor the $(\text{NM} - ga)_k$ have a dual when $k > 0$. The case $k = 0$ is referred to as the multivariate negative binomial case. The case $k = 0$ has a dual if and only if $n = 2$ or $n = 3$.

Now it is shown that there is no dual for the multivariate negative binomial distribution. Denote by $\mathbb{N}$ the set of nonnegative integers. Consider the measure on $\mathbb{N}^n$ defined by $\mu(dx) = (\delta_0 - \delta_{e_1} - \cdots - \delta_{e_n})^{-1} *$ where the inverse is taken in the sense of convolution. Its Laplace transform is

$$B_{\mu}(s) = \frac{1}{1 - e^{-s_1} - \cdots - e^{-s_n}} = \frac{1}{\Delta},$$

defined on the set

$$S(\mu) = \{ s \in \mathbb{R}^n; \Delta > 0 \}.$$

We show that no dual $\mu^*$ of $\mu$ exists. Suppose the existence of $\mu^*$. Since $\ell_\mu = -\log \Delta$ we have

$$m = (m_1, \ldots, m_n) = -\ell_\mu'(s) = \frac{1}{\Delta}(e^{-s_1}, \ldots, e^{-s_n}).$$

Since $e^{ni} = m_i \Delta$ we get $\Delta(m_1 + \cdots + m_n) = e^{-s_1} + \cdots + e^{-s_n} = 1 - \Delta$ leading to $\Delta = 1/(1 + m_1 + \cdots + m_n)$. Finally we have on the domain $m_i \geq 0$, for all $i \in \{1, \ldots, n\}$,

$$\ell_\mu'(m) = (-s_1, \ldots, -s_n)$$

with $-s_i = \log m_i - \log(1 + m_1 + \cdots + m_n)$ and therefore

$$\ell_\mu'(m) = m_1 \log m_1 + \cdots + m_n \log m_n - (1 + m_1 + \cdots + m_n) \log(1 + m_1 + \cdots + m_n),$$

$$B_{\mu'}(m) = \frac{m_1^{m_1} \cdots m_n^{m_n}}{(1 + m_1 + \cdots + m_n)^{1 + m_1 + \cdots + m_n}}$$

(33)

For fixed $m$ and $m'$ the matrix

$$M = \begin{bmatrix} B_{\mu'}(2m) & B_{\mu'}(m + m') \\ B_{\mu'}(m + m') & B_{\mu'}(2m') \end{bmatrix}$$

(34)

is semi positive definite since

$$(a, b)M \begin{pmatrix} a \\ b \end{pmatrix} = a^2 B_{\mu'}(2m) + 2ab B_{\mu'}(m + m') + b^2 B_{\mu'}(2m') = \int_{\mathbb{R}^n} (ae^{-\langle m, x \rangle} + be^{-\langle m', x \rangle}) \mu^*(dx) \geq 0.$$

This implies that $\det M \geq 0$. Let us apply this remark to the particular case $m = (0, 1, 0, \ldots, 0)$ and $m' = (1, 1, 0, \ldots, 0)$. It leads to

$$\det M = \frac{1}{3^4 4^5 5^3} (4^8 - 3^2 5^3) = -\frac{19339}{3^4 4^5 5^3} < 0$$

which is the desired contradiction.

Next it is shown that there is no dual for the $(\text{NM} - ga)_{n-1}$ case. We use the notation of [14]. This exponential family is generated by the measure $\mu(dx_1, \ldots, dx_{n-1}, dy)$ on $\mathbb{N}^{n-1} \times (0, \infty)$ defined by

$$(\delta_0 - \delta_{e_1} - \cdots - \delta_{e_{n-1}})^{-1}*(dx_1, \ldots, dx_{n-1}) \frac{y^{x_1 + \cdots + x_{n-1}}}{(x_1 + \cdots + x_{n-1})!} 1_{(0,\infty)}(y)dy.$$
The exponential family $F(\mu)$ is the set of the laws of $(X, Y)$ when $X = (X_1, \ldots, X_{n-1})$ is multivariate negative binomial on $\mathbb{N}^{n-1}$ and the conditional law of $Y$ knowing $X$ is gamma with shape parameter $1 + |X| = 1 + X_1 + \cdots + X_{n-1}$. Its variance function is

$$V_\mu(m) = m \otimes m + \text{diag}(m_1, \ldots, m_{n-1}, 0)$$

and the Laplace transform of $\mu$ is

$$B_\mu(s) = (s_0 - e^{-s_1} - \cdots - e^{-s_{n-1}})^{-1} = \frac{1}{\Delta}.$$ 

We show first the following: If $\mu^*$ exists then its Laplace transform is defined on $m_i > 0$, $i \in \{1, \cdots, n\}$, by

$$B_{\mu^*}(m) = \frac{m_1^{m_1} \cdots m_{n-1}^{m_{n-1}}}{m_{n-1}^{m_{n-1}+\cdots+m_{n-1}}}.$$  

(35)

To see this, since $\Delta' = (e^{-s_1}, \ldots, e^{-s_{n-1}}, 1)$ and that $m = -\ell_{\mu^*}'(s) = \Delta' / \Delta$ one gets $m_n = 1 / \Delta$ and $e^{-s_i} = m_i / m_n$ for $i < n$ leading to

$$\ell_{\mu^*}'(m) = -s = \log(m_1 / m_n), \ldots, \log(m_{n-1} / m_n), (1 + m_1 + \cdots + m_{n-1}) / m_n$$

from which the result is obtained.

To see that actually $\mu^*$ does not exist we use the same method as for the multivariate negative binomial distribution by observing that the matrix $M$ defined by (34) has a negative determinant when choosing $m = (0, \ldots, 0, 1)$ and $m' = (1, \ldots, 0, 1).

Now it will be established that there is no dual for the $(\text{NM - ga})_k$ case with $0 < k < n - 1$. This exponential family is generated by the measure $\mu(dx_1, \ldots, dx_3, dy, dz)$ on $\mathbb{N}^k \times (0, \infty) \times \mathbb{R}^{n-k-1}$ defined by

$$(\delta_0 - \delta_{\nu_1} - \cdots - \delta_{\nu_k})^{-1}(dx_1, \ldots, dx_3) \frac{y^{x_1+\cdots+x_3}}{(x_1 + \cdots + x_3)^{y_{\nu_1}+\cdots+y_{\nu_k}}}(0,0,0)(y)dy \times N(0, yI_{n-k-1})(dz),$$

where $N(0, yI_{n-k-1})$ is the Gaussian distribution on $\mathbb{R}^{n-k}$ with covariance matrix $yI_{n-k-1}$. Its Laplace transform is $B_\mu(s) = 1 / \Delta$ with

$$\Delta = s_{k+1} - \frac{1}{2} \sum_{j=k+2}^n s_j^2 - \sum_{i=1}^k e^{-s_i}.$$

Computations, quite analogous to the previous one, yield the Laplace transform of $\mu^*$ would be

$$B_{\mu^*}(m) = \frac{\prod_{j=1}^k e^{-m_j} m_j^{m_j}}{m_{k+1}^{m_{k+1}}} \exp \left( \frac{1}{2} \sum_{j=k+2}^n \frac{m_j^2}{m_{k+1}} \right).$$  

(36)

If $k \neq 0$, putting $m_{k+2} = \cdots = m_n = 0$ shows that $B_{\mu^*}(m)$ cannot be a Laplace transform as we have seen for (35).

Consider now the $(\text{NM - ga})_0$ case $B_\mu(s) = 1 / \Delta$ with

$$\Delta = s_1 - \frac{1}{2} \sum_{j=2}^n s_j^2.$$

This generates the exponential family containing the distribution of $(X_1, Z_2 \sqrt{X_1}, \ldots, Z_n \sqrt{X_1})$, where $X_1, Z_2, \ldots, Z_n$ are independent, with $Z_i \sim N(0, 1)$ and where $X_1$ is exponential with mean 1. Standard calculations show that if $\mu^*$ exists then

$$B_{\mu^*}(m) = \frac{1}{m_1} \exp \left( \frac{1}{2} \sum_{j=2}^n \frac{m_j^2}{m_1} \right).$$

Denoting for simplicity $\bar{m} = (m_2, \ldots, m_n)$ we can write

$$B_{\mu^*}(m_1, \bar{m}) = \int_{\mathbb{R}^{n-1}} e^{-((\bar{m}, \bar{s}) - \frac{1}{2} m_1 \bar{s})^2} m_1^{(n-3)/2} m_1^{(n-3)/2} \frac{dx}{(2\pi)^{(n-1)/2}}$$

If $n - 3 \leq 0$ we prove that $B_{\mu^*}$ is a Laplace transform as follows:
If \( n = 2 \) we use

\[
\frac{1}{\sqrt{m_1}} = \int_0^\infty e^{-m_1 t} \frac{dt}{\sqrt{\pi t}}
\]

for writing

\[
B_{\mu'}(m_1, m_2) = \frac{1}{\sqrt{2\pi}} \int e^{-m_2 s} \int_0^\infty e^{-m_1(t+\frac{1}{2} s^2)} dt \frac{dx}{\sqrt{\pi t}} = \frac{1}{\sqrt{2\pi}} \int e^{-m_2 s} \int_{x^2/2}^\infty e^{-m_1 y} dy \frac{dx}{\sqrt{\pi(y - \frac{1}{2} x^2)}}.
\]

This shows that \( B_{\mu'} \) is the Laplace transform of the density \( \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\pi(y - \frac{1}{2} x^2)}} \) restricted to the interior of the parabola \( y = x^2/2 \).

If \( n = 3 \) we have

\[
B_{\mu'}(m_1, \bar{m}) = \frac{1}{\pi} \int e^{-(\|x\|^2 - \frac{1}{2} m_1 x_1^2)} dx.
\]

This is the Laplace transform of the following singular measure in \( \mathbb{R}^3 \): it is concentrated on the cone \( y = \frac{1}{2} (x_1^2 + x_2^2) \) and it is the image of the Lebesgue measure \( \frac{dx_1 dx_2 dx_3}{\sqrt{\pi}} \) on \( \mathbb{R}^3 \) by the map \( x \mapsto (x, \frac{1}{\sqrt{2}} \|x\|^2) \).

To complete the study of these last two measures \( \mu' \), one can describe the variance function of \( F(\mu') \) with the following matrices for \( n = 2 \) and \( n = 3 \):

\[
V_{F(\mu')}(s_1, s_2) = \left( s_1 - \frac{1}{2} s_3^2 \right) \begin{bmatrix}
    s_1 + \frac{1}{2} s_2^2 & \frac{1}{2} s_2 & 1 \\
    \frac{1}{2} s_2 & 1 & 0 \\
    1 & 0 & 1
\end{bmatrix},
\]

\[
V_{F(\mu')}(s_1, s_2, s_3) = \left( s_1 - \frac{1}{2} (s_2^2 + s_3^2) \right) \begin{bmatrix}
    s_1 + \frac{1}{2} (s_2^2 + s_3^2) & \frac{1}{2} s_2 & \frac{1}{2} s_2 & \frac{1}{2} s_3 \\
    \frac{1}{2} s_2 & 1 & 0 & 1 \\
    \frac{1}{2} s_2 & 1 & 0 & 1 \\
    \frac{1}{2} s_3 & 0 & 1 & 0
\end{bmatrix}.
\]

We skip this standard computation.

For \( n \geq 4 \) one can suspect that \( \mu' \) does not exist. Enough is to prove it for \( n = 4 \), since we can do \( m_5 = \ldots = m_n = 0 \) to pass from the case \( n > 4 \) to the case \( n = 4 \). For simplicity of calculations in the case \( n = 4 \) we rather set \( m_1 = t \) and \( m_2 = s_1, m_3 = s_2, m_4 = s_3 \). And we show now that

\[
L(\vec{s}, t) = \frac{1}{t + 1} \exp \left( -\frac{1}{2(t+1)} \|\vec{s}\|^2 \right)
\]

is not a Laplace transform on \((-1, \infty) \times \mathbb{R}^3 \). Suppose the contrary. Consider a random variable \((\vec{X}, Y)\) such that for \( t > -1 \)

\[
A = e^{-(\|\vec{X}\|^2 - Y)}, \quad \mathbb{E}(A) = L(\vec{s}, t),
\]

Denote \( \vec{1} = (1, 1, 1) \). We are going to prove that for a suitable \((\vec{r}, \vec{s})\) we have

\[
\mathbb{E}(\|\vec{X} - r\vec{Y}\|^2 A) < 0
\]

(37)

which proves the impossibility. By taking partial derivatives of \( \mathbb{E}(\|\vec{X} - rY\|^2 A) \) a calculation gives \( \mathbb{E}(\|\vec{X} - r\vec{Y}\|^2 A) = \frac{1}{(t+1)^2} e^{\|\vec{s}\|^2} B \)

with the shorter notations \( u = \frac{1}{2} \|\vec{s}\|^2 \) and

\[
B = 2u(t+1)^2 + 2r(\vec{s}, \vec{1})((t+1)^2 + (t+1)^2(1 + u)) + r^2(2(t+1)^2 + 4u(t+1) + u^2)
\]

We now choose \( \vec{s} = -2r\vec{1} \). This gives \( u = 6r^2 \) and \( (s_1, s_2, s_3) = 6r \). The quantity \( B \) becomes

\[
B = -12r(t+1)^2 + (t+1)^2(1 + 6r^2) + r^2(14(t+1)^2 + 24(t+1)r^2 + 36r^4).
\]

For fixed \( t \) we can now choose \( r \) small enough such that \( B < 0 \). Therefore (37) is proved.

Finally it is shown that there is no dual for the multivariate hyperbolic case. Following Casalis [4], p. 1836, the multivariate hyperbolic case is described by the Laplace transform

\[
B_\mu(s) = (\cos s_n - e^{-s_1} - \ldots - e^{s_{n-1}})^{-1}
\]

23
defined on
\[ S(\mu) = \{ s \in \mathbb{R}^{n-1} \times \left( -\frac{1}{2}, \frac{1}{2} \right) : \cos s_1 - e^{-s_1} - \cdots - e^{-s_{n-1}} > 0 \}. \]

We spare to the reader the proof of the fact that if a dual \( \mu^* \) exists then its Laplace transform is, with the notation
\[ S = 1 + m_1 + \cdots + m_{n-1}, \]
\[ B_{\mu^*}(m) = \frac{\prod_{i=1}^{n-1} m_i^{m_i}}{(S^2 + m_2^{1/2})^{n/2}} \exp(m_n \arctan(m_n/S)). \]

Note that if \( m_1 = \cdots = m_{n-1} = 0 \) we get \( B_{\mu^*}(m) = h_2(m_n) \) where \( h_2 \) is defined in \([13]\). But is has been proven in Section 3 that \( h_2 \) cannot be a Laplace transform. As a consequence the present \( \mu^* \) does not exist.

8. Remarks.

Duality, with its links with the large deviations as described in Proposition\([5]\) has not a strong probabilistic interpretation compared to the notion of reciprocity, which was used in Letac and Mora\([15]\). However it widens the zoo of variance functions, and it creates unexpected links like the pair Poisson-Landau. The problem of deciding of the do some measures have dual and why others have not? The Babel class has been defined and described in Letac\([16]\), existence of a dual is unsolved in many circumstances, and we need more tools than we described in Section 3.5. Why would also be in order. The role of steepness is not well understood: all tractable examples of a non steep \( \mu \) lead to the non existence of a dual \( \mu^* \). Does non steepness imply no dual? Dealing with exponential families in \( \mathbb{R}^n \) is generally hard: the simplest non trivial diagonal family in \( \mathbb{R}^2 \) is generated by the measure \( \delta_0 + \delta_1 + \delta_c + c\delta_{1,2r} \) where \( c \neq 1 \) (see\([2]\) p. 895) and leads to computations of great complexity. We have not tried yet the cubic families in \( \mathbb{R}^n \) for \( n > 1 \) characterized by Hassairi\([8]\).

Acknowledgments

I have had some discussions with Shaul Bar-Lev, Lev Klebanov and Vladimir Vinogradov about various parts of the paper. I thank them all.

References

[1] S.K. Bar-Lev, P. Enis, Reproducibility and natural exponential families with power variance functions, Ann. Statist. 14 (1986) 1507–1522.
[2] S.K. Bar-Lev, D. Bshouty, P. Enis, G. Letac, I. Liu Lu, D. Richards, The diagonal multivariate natural exponential families and their classifica-
tion, J. Theoret. Probab. 7 (1994) 883–929.
[3] M. Casalis, Les families exponentielles à variance quadratique homogène sont des lois de Wishart sur un cône symétrique, C.R. Acad. Sci.
Paris. Sér. I Math. 312 (1991) 143–146.
[4] M. Casalis, The 2d + 4 simple quadratic natural exponential families on \( \mathbb{R}^2 \), Ann. Statist. 24 (1996) 1828–1864.
[5] M. Casalis, G. Letac, The Lukacs-Olkin-Rubin characterization of the Wishart distribution on symmetric cones, Ann. Statist. 24 (1996) 763–786.
[6] H. Cramér, Sur un nouveau théorème-limite de la théorie des probabilités, Actualités Scientifiques et Industrielles, 736 (1938) 5–23.
[7] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II, Wiley, New York, 1966.
[8] A. Hassairi, La classification des familles exponentielles de \( \mathbb{R}^n \) par l’action du groupe linéaire de \( \mathbb{R}^{n+1} \), C.R. Acad. Sci. Paris. Sér. I Math. 315 (1992) 207–210.
[9] B. Jørgensen, Exponential dispersion models, J. Roy. Statist. Soc. Ser. B 49 (1987) 127–162.
[10] B. Jørgensen, The Theory of Dispersion Models, Chapman and Hall, London, 1997.
[11] T. Kawata, Fourier Analysis in Probability Theory, Academic Press, New York, 1972.
[12] D. Kirilov, Dilogarithm identities, Progress of Theoretical Physics Supplement 118 (1994) 61–142 arXiv 9408113v2.
[13] C.C. Kokonendji, Exponential families with variance functions in \( \sqrt[3]{\mathbb{R}} \): Seshadri’s class, Test 3 (1994) 123–172.
[14] G. Letac, Le problème de la classification des familles exponentielles naturelles de \( \mathbb{R}^2 \) ayant une fonction variance quadratique, Probability on Groups IX, Oberwolfach, Lecture notes in Mathematics, Springer 1379 (1989) 192–216.
[15] G. Letac, M. Mora, Natural exponential families with cubic variances, Ann. Statist. 18 (1990) 1–37.
[16] G. Letac, Lectures on Natural Exponential Families and Their Variance Functions, 50, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1992.

[17] G. Letac, Integration and Probability, Exercises and Solutions, Springer, New York, 1995.

[18] G. Letac, Associated exponential families and elliptic functions, in The Fascination of Probability, Statistics and their Applications in honor of Ole Barndorff-Nielsen, eds. M. Podolskij, R. Stelzer, S. Thorbjørnsen and A. Veraart pp. 53–84 2016, Springer, New York.

[19] L. Lewin, Polylogarithms and Associated Functions, North Holland, New York, 1981.

[20] M. Marucho, C. Garcia-Canal, H. Fanchiotti, The Landau distribution for charged particles traversing thin films, Internat. J. Modern Phys. C 17 (2006) 1461-1476 arXiv hep-ph/0305310v3.

[21] C. Morris, Natural exponential families with quadratic variance functions, Ann. Statist. 10 (1982) 65–80.

[22] M. Tweedie, An index which distinguishes between some important exponential families, in Statistics: Applications and New Directions. Proc. Indian Institute Golden Jubilee Internat. Conf., eds. J.K. Ghosh and J. Roy pp. 579–604 1984, Indian Statistical Institute, Calcutta.

[23] V. Vinogradov, R. Paris, On two extensions of the canonical Feller-Spitzer distribution, J. Stat. Dist. Appl. 8 (2021) 1–25.

[24] D. Zagier, The Dilogarithm Function, in Frontiers in Number Theory, Physics, and Geometry, Vol. II, eds. P. Cartier, P. Moussa, B. Julia and P. Vanhove pp. 3–65 2007, ISBN 978-3-540-30308-4 DOI 10.1007/978-3-540-30308-4 -1, on line.