Discussiones Mathematicae
Probability and Statistics 34 (2014) 35–49
doi:10.7151/dmps.1167

ON THE PROPERTIES OF THE GENERALIZED NORMAL DISTRIBUTION

Thomas L. Toulias and Christos P. Kitsos

Technological Educational Institute of Athens
Informatics Department
Egaleo 12243, Athens, Greece

e-mail: {t.toulias, xkitsos}@teiath.gr

Abstract

The target of this paper is to provide a critical review and to enlarge the theory related to the Generalized Normal Distributions (GND). This three term (position, scale shape) distribution is based in a strong theoretical background due to Logarithm Sobolev Inequalities. Moreover, the GND is the appropriate one to support the Generalized entropy type Fisher's information measure.

Keywords: entropy type Fisher's information, Shannon entropy, Normal distribution, truncated distribution.

2010 Mathematics Subject Classification: 60E05, 62H10, 62E15.

1. Introduction

The Normal (Gaussian) distribution is the most important tool of Statistics for both the theoretical framework and the application field sice the time that Gauss offered it. In this section we briefly review the extensions development.

Let X be a random variable (r.v.) from the Normal distribution with mean μ and variance $\sigma^2 > 0$, $X \sim \mathcal{N}(\mu, \sigma^2)$. Then it is well known that the probability density function (p.d.f.) of X is of the form

$$\phi_1(x;\ \mu, \sigma) = C_2^1(\sigma) \exp\left\{-\tfrac{1}{2} Q_\theta(x)\right\}, \quad x \in \mathbb{R}, \tag{1}$$

with $C_2^1(\sigma) = 1/(\sqrt{2\pi}\sigma)$ a constant (normalizing) factor, depending on σ and Q_θ is the quadratic form

$$Q_\theta(x) = \tfrac{1}{\sigma^2}(x-\mu)^2, \quad x \in \mathbb{R},$$

while the parameter $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+^*$.

The computational tractability of the Normal distribution can be widely expanded. For the pair of r.v.'s $(X_1, X_2) = X$ the bivariate Normal distribution $\mathcal{N}^2(\mu, \Sigma)$ with $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$ and $\Sigma \in \mathbb{R}^{2\times 2}$, such that $\Sigma = (\Sigma_{ij})$, $\Sigma_{ii} = \sigma_i^2$, $\Sigma_{ij} = \rho\sigma_i\sigma_j$, $i, j = 1, 2$, is defined through the p.d.f.

$$\phi_2(x_1, x_2) = C_2^2(\sigma_1, \sigma_2)\exp\{-\tfrac{1}{2}Q_\theta(x_1, x_2)\}, \quad (x_1, x_2) \in \mathbb{R}^2. \tag{2}$$

The constant (normalizing) factor $C_2^2(\sigma_1, \sigma_2)$ depends on the scales of X_1 and X_2, i.e., σ_1, σ_2, and the correlation $\rho = \mathrm{Corr}(\sigma_1, \sigma_2)$, namely

$$C_2^2(\sigma_1, \sigma_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} = \frac{1}{2\pi\sqrt{|\Sigma|}}, \tag{3}$$

$Q_\theta(x_1, x_2)$, $\theta = (\mu, \Sigma)$, is the quadratic form

$$Q_\theta(x_1, x_2) = \tfrac{1}{2(1-\rho^2)}\left(z_1^2 - 2\rho z_1 z_2 + z_2^2\right), \quad (x_1, x_2) \in \mathbb{R}^2, \tag{4}$$

where $z_i = (x_1 - \mu_1)/\sigma_i$, $i = 1, 2$. It is essential that the marginal distributions of X are $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $i = 1, 2$, while the inverse is not true. The case $\rho = 0$ is equivalent to the independence of X_1 and X_2, while with ρ close to zero the conditional variances of $X_1|X_2 = x_2$ and $X_2|X_1 = x_1$ are closed to the variances σ_1 and σ_2 respectively.

The bivariate Normal distribution is an intermediate extension for the p-variate r.v. X, i.e., for the multivariate Normal $\mathcal{N}^p(\mu, \Sigma)$, with $\mu \in \mathbb{R}^p$ and $\Sigma \in \mathbb{R}^{p\times p}$ with dominating measure the Lebesque measure on $\mathbb{R}^p$ with density

$$\phi_p(x;\ \mu, \Sigma) = C_2^p(\Sigma)\exp\left\{-\tfrac{1}{2}Q_\theta(x)\right\}, \quad x \in \mathbb{R}^p, \tag{5}$$

normalizing factor

$$C_2^p(\Sigma) = (2\pi)^{-p}|\Sigma|^{-1/2}, \tag{6}$$

and Q_θ the p-quadratic form

$$Q_\theta(x) := (x-\mu)\Sigma^{-1}(x-\mu)^{\mathrm{T}},$$

with $\theta := (\mu, \Sigma) \in \mathbb{R}^{p\times p\times p}$. Notice that for the r.v. X from the p-variate Normal and A a given $p \times p$ matrix, it holds

$$X \sim \mathcal{N}^p(\mu, \Sigma) \Rightarrow \mathrm{A}X \sim \mathcal{N}^p(\mathrm{A}\mu, \mathrm{A}\Sigma\mathrm{A}^{\mathrm{T}}). \tag{7}$$

Following the above discussion the γ-order GND for the p-variate r.v. X, $\mathcal{N}^p_\gamma(\mu, \Sigma)$ say, was defined as an extremal of (an Euclidean) Logarithm Sobolev Inequality (LSI). Following [8] adopting Gross Logarithm Inequality [7], with respect to the Gaussian weight, holds

$$\int\limits_{\mathbb{R}^p} \|g\|^2 \log \|g\|^2 dm \leq \tfrac{1}{\pi} \int\limits_{\mathbb{R}^p} \|\nabla g\|^2 dm, \tag{8}$$

where $\|g\|_2 = 1$, $dm = \exp\{-\pi|x|^2\}dx$ ($\|g\|_2$ is the norm in $\mathscr{L}^2(\mathbb{R}^p, dm)$). Inequality (8) is equivalent to the (Euclidean) LSI,

$$\int\limits_{\mathbb{R}^p} \|u\|^2 \log \|u\|^2 dx \leq \tfrac{p}{2} \log \left\{ \tfrac{2}{\pi p e} \int\limits_{\mathbb{R}^p} \|\nabla u\|^2 dx \right\}, \tag{9}$$

for any function $u \in \mathscr{W}^{1,2}(\mathbb{R}^p)$ with $\int_{\mathbb{R}^p} |u|^2 dx = 1$, see [8] for details. This inequality is is optimal with extremals $u(x)$ Gaussians. Now, consider the extension of Del Pino and Dolbeault in [20] for the LSI as in (9). For any $u \in \mathscr{W}^{1,2}(\mathbb{R}^p)$ with $\|u\|_\gamma = 1$, the γ-LSI holds, i.e.,

$$\int\limits_{\mathbb{R}^p} \|u\|^\gamma \log \|u\|^\gamma dx \leq \tfrac{p}{\gamma} \log \left\{ K_\gamma \int\limits_{\mathbb{R}^p} \|\nabla u\|^\gamma dx \right\}, \tag{10}$$

with the optimal constant K_γ equals to

$$K_\gamma = \tfrac{\gamma}{p}(\tfrac{\gamma-1}{e})^{\gamma-1}\pi^{-\gamma/2}(\xi^p_\gamma)^{\gamma/p}, \tag{11}$$

where

$$\xi^p_\gamma = \frac{\Gamma(\frac{p}{2}+1)}{\Gamma(p\frac{\gamma-1}{\gamma}+1)}, \tag{12}$$

and $\Gamma(\cdot)$ the usual gamma function.

Inequality (10) is optimal and the equality holds when a new "hyper" multivariate Normal distribution with mean vector $\mu \in \mathbb{R}^{p\times 1}$, scale matrix $\Sigma \in \mathbb{R}^p$ and a new (shape) parameter $\gamma \in \mathbb{R}\setminus[0,1]$ is considered of the form $u(x) = f_\gamma(x)$, $x \in \mathbb{R}$ as

$$f_\gamma(x;\ \mu, \Sigma) = C^p_\gamma(\sigma) \exp\{-\tfrac{\gamma-1}{\gamma}[Q_\theta(x)]^{\frac{\gamma}{2(\gamma-1)}}\}, \quad x \in \mathbb{R}^p, \tag{13}$$

with normalizing factor

$$C_2^p(\Sigma) = \pi^{-p}|\Sigma|^{-1/2}\xi_\gamma^p(\tfrac{\gamma-1}{\gamma})^{p\frac{\gamma-1}{\gamma}}, \tag{14}$$

and p-quadratic form $Q_\theta(x) := (x-\mu)\Sigma^{-1}(x-\mu)^{\mathrm{T}}$, where $\theta := (\mu,\Sigma) \in \mathbb{R}^{p\times p\times p}$. We shall refer to the above distribution as the *generalized γ-order Normal distribution*, or γ-GND. Notice that with $\gamma = 2$ the γ-order Normal $\mathcal{N}_\gamma^P(\mu,\Sigma)$ is reduced to to the usual multivariate Normal $\mathcal{N}^P(\mu,\Sigma)$. The elliptical contoured γ-GND is reduced to spherical contoured when $\Sigma = \sigma^2\mathbb{I}_p$. An immediate result is that the maximum density value $f_\gamma(\mu) = C_\gamma^p$. Recall that $\phi_1(x) = C_2^1 = 1/(\sqrt{2\pi}\sigma)$.

One of the merits of the γ-GND defined above is that belongs to the Kotz type distributions family, [15].

In Section 2 we provide a compact critical review of the properties of the generalized γ-order Normal distribution considered also as a generator for other distributions. In Section 3 information-theoretic results are obtained through the generalized entropy type information measure, while in Section 4 a discussion is provided.

2. On the family of the γ-GND

Recall the second-ordered Normal $\mathcal{N}_2^p$, i.e., the multivariate Normal distribution $\mathcal{N}^p(\mu,\Sigma)$, the multivariate Uniform $\mathcal{U}^p(\mu,\Sigma)$, Laplace $\mathcal{L}^p(\mu,\Sigma)$, as well as the degenerate Dirac distribution $\mathcal{D}^p(\mu)$ with p.d.f. $f_\mathcal{U}$, $f_\mathcal{L}$, $f_\mathcal{D}$ as follows:

$$f_\mathcal{U}(x) = \frac{\Gamma(\frac{p}{2}+1)}{(\pi^p \det\Sigma)^{1/2}}, \quad x\in\mathbb{R}^p, \quad \text{with} \quad Q_\theta(x)\le 1, \tag{15}$$

$$f_\mathcal{L}(x) = \frac{\Gamma(\frac{p}{2}+1)}{p!(\pi^p \det\Sigma)^{1/2}} \exp\left\{-Q_\theta^{1/2}(x)\right\}, \quad x\in\mathbb{R}^p, \tag{16}$$

$$f_\mathcal{D}(x) = \begin{cases} +\infty, & x=\mu, \\ 0, & x\in\mathbb{R}^p\setminus\mu. \end{cases} \tag{17}$$

The following Theorem states that the above distributions are members of the γ-GND family for certain values of the shape parameter γ. This also provides evidence that the order γ is essential as, eventually, "bridges" distributions with complete different shape "attitude", see Figure 1, as well as "heavy-tailness", see [4] for an economical example.

Let X be a random variable following $\mathcal{N}_\gamma^p(\mu,\sigma^2\mathbb{I}_p)$. Then the "probability mass" around μ of X with radius $\sigma = 1$, i.e., $\Pr\{\|X-\mu\|\le 1\} = \int_{\|X-\mu\|\le 1} f_(x)dx$ for dimensions $p = 1,2,3$ has been evaluated in Table 1 below.

Table 1. Probability mass values for various $X \sim \mathcal{N}_\gamma(\mu, \mathbb{I}_p)$, $p = 1, 2, 3$.

γ	$\Pr\{\lvert X-\mu\rvert \le 1\}$ **p = 1**	$\Pr\{\lVert X-\mu\rVert \le 1\}$ **p = 2**	$\Pr\{\lVert X-\mu\rVert \le 1\}$ **p = 3**
−100	0.6315	0.8633	0.9491
−10	0.6262	0.8516	0.9392
−2	0.6084	0.8100	0.8995
−1	0.5940	0.7737	0.8603
−0.05	0.5290	0.5889	0.6233
1	**1.0000**	**1.0000**	**1.0000**
2	**0.6827**	**0.9545**	**0.9973**
5	0.6470	0.8953	0.9724
10	0.6390	0.8792	0.9614
100	0.6328	0.8669	0.9513
$\pm\infty$	**0.6320**	**0.8660**	**0.9510**

Theorem 1. *The multivariate γ-GND $\mathcal{N}_\gamma^p(\mu, \Sigma)$ with p.d.f. f_γ, coincides for different values of γ with the Uniform, Normal, Laplace and Dirac distributions in terms that*

$$f_\gamma = \begin{cases} f_{\mathcal{D}}, & \text{for } \gamma = 0 \text{ and } p = 1, 2, \\ 0, & \text{for } \gamma = 0 \text{ and } p \ge 3, \\ f_{\mathcal{U}}, & \text{for } \gamma = 1, \\ f_{\mathcal{N}}, & \text{for } \gamma = 2, \\ f_{\mathcal{L}}, & \text{for } \gamma = \pm\infty. \end{cases} \tag{18}$$

See [11].

The linear relation described in (7) for the multivariate Normal is valid for the γ-GND, in the sense that for given A an appropriate matrix and b an appropriate vector, then

$$X \sim \mathcal{N}_\gamma^p(\mu, \Sigma) \Rightarrow \mathrm{A}X + b \sim \mathcal{N}_\gamma^p(\mathrm{A}\mu + b, \mathrm{A}\Sigma\mathrm{A}^{\mathrm{T}}). \tag{19}$$

Simple calculation also proves that if the matrix A is reduced to an appropriate vector, relation (19) is still valid.

Recall that for the univariate Normal distribution (2-GND) the cumulative distribution function (c.d.f.) for the standardized normally distributed $Z \sim \mathcal{N}(0, 1)$ is

$$\Phi(z) = \tfrac{1}{2} + \tfrac{1}{2}\operatorname{erf}(\tfrac{z}{2}), \quad z \in \mathbb{R}, \tag{20}$$

with $\operatorname{erf}(\cdot)$ being the usual error function. For the γ-GND the generalized error function Erf_a is involved [6]. Indeed the $\mathrm{Erf}_{\gamma/(\gamma-1)}$ is considered and the following holds.

Theorem 2. *Let X be a random variable from the univariate γ-GND, i.e., $X \sim \mathcal{N}_\gamma^p(\mu, \sigma^2)$ with p.d.f. f_γ. If F_γ is the c.d.f. of f_γ and Φ_γ the c.d.f. of the standardized $Z = \frac{1}{\sigma}(X-\mu) \sim \mathcal{N}_\gamma(0,1)$, then*

$$(21)\quad F_\gamma(x) = \Phi_\gamma(\tfrac{x-\mu}{\sigma}) = \tfrac{1}{2} + \frac{\sqrt{\pi}}{2\,\Gamma(\frac{\gamma-1}{\gamma})\,\Gamma(\frac{\gamma}{\gamma-1})}\operatorname{Erf}_{\frac{\gamma}{\gamma-1}}\left\{(\tfrac{\gamma-1}{\gamma})^{\frac{\gamma-1}{\gamma}}\tfrac{x-\mu}{\sigma}\right\}, \quad x \in \mathbb{R}.$$

Proof. We have

$$F_\gamma(x) = \int_0^x f_\gamma(t)dt = \frac{C_\gamma^1}{\sigma}\int_{-\infty}^x \exp\left\{-\tfrac{\gamma-1}{\gamma}\left|\tfrac{x-\mu}{\sigma}\right|^{\frac{\gamma}{\gamma-1}}\right\}dt.$$

Applying the linear transformation $w = \frac{\log t-\mu}{\sigma}$, the above is reduced to

$$(22)\qquad F_{X_\gamma}(x) = C_\gamma^1 \int_{-\infty}^{\frac{1}{\sigma}(x-\mu)} \exp\{-\tfrac{\gamma-1}{\gamma}|w|^{\frac{\gamma}{\gamma-1}}\}dw = \Phi_{Z_\gamma}(\tfrac{x-\mu}{\sigma}),$$

where Φ_{Z_γ} is the c.d.f. of the standardized γ-ordered Normal distribution with $Z_\gamma = \frac{1}{\sigma}(X_\gamma - \mu) \sim \mathcal{N}_\gamma(0,1)$. Moreover, Φ_{Z_γ} can be expressed in terms of the generalized error function. In particular

$$\Phi_{Z_\gamma}(z) = C_\gamma^1\int_{-\infty}^z \exp\{-\tfrac{\gamma-1}{\gamma}|w|^{\frac{\gamma}{\gamma-1}}\}dw = \Phi_{Z_\gamma}(0) + C_\gamma^1\int_0^z \exp\{-\tfrac{\gamma-1}{\gamma}|w|^{\frac{\gamma}{\gamma-1}}\}dw,$$

and as f_{Z_γ} is a symmetric density function around zero, we have

$$\Phi_{Z_\gamma}(z) = \exp\left\{-\tfrac{\gamma-1}{\gamma}|w|^{\frac{\gamma}{\gamma-1}}\right\}dw = \tfrac{1}{2} + C_\gamma^1\int_0^z \exp\left\{-\left|(\tfrac{\gamma-1}{\gamma})^{\frac{\gamma-1}{\gamma}}w\right|^{\frac{\gamma}{\gamma-1}}\right\}dw,$$

and thus

$$(23)\qquad \Phi_{Z_\gamma}(z) = \tfrac{1}{2} + C_\gamma^1(\tfrac{\gamma}{\gamma-1})^{\frac{\gamma-1}{\gamma}}\int_0^{(\frac{\gamma-1}{\gamma})^{\frac{\gamma-1}{\gamma}}z} \exp\left\{-u^{\frac{\gamma}{\gamma-1}}\right\}du.$$

Substituting the normalizing factor, as in (14), we obtain

$$(24)\qquad \Phi_{Z_\gamma}(z) = \tfrac{1}{2} + \frac{\sqrt{\pi}}{2\,\Gamma(\frac{\gamma-1}{\gamma}+1)\,\Gamma(\frac{2\gamma-1}{\gamma-1})}\operatorname{Erf}_{\frac{\gamma}{\gamma-1}}\left\{(\tfrac{\gamma-1}{\gamma})^{\frac{\gamma-1}{\gamma}}z\right\}, \quad z \in \mathbb{R},$$

through the definition of the generalized error function, i.e., (21) holds. ∎

It is interesting to notice that for $X \sim \mathcal{N}_\gamma^p(\mu, \Sigma)$ then $\mathrm{E}(X) = \mu$ and

$$\operatorname{Cov} X = \frac{\Gamma\left((p+2)\frac{\gamma-1}{\gamma}\right)}{\Gamma(p\frac{\gamma-1}{\gamma})} (\tfrac{\gamma}{\gamma-1})^{2\frac{\gamma-1}{\gamma}} (\operatorname{rank}\Sigma)^{-1}\Sigma,$$

see [11], and therefore, for the usual Normal case of the 2-GND, the scale matrix Σ is indeed the covariance, i.e., $\Sigma = \operatorname{Cov} X$ with $\gamma = 2$.

As far as the characteristic function $h_\gamma(t)$, $t \in \mathbb{R}^p$ is concerned for the positive-ordered Normal r.v. $X \sim \mathcal{N}_\gamma^p$ with $\gamma > 1$, the following Theorem holds, see [11] for details.

Theorem 3. *Let h_γ being the characteristic function of the multivariate γ-GND r.v. $X \sim \mathcal{N}_\gamma^p(\mu, \Sigma)$ with $\gamma > 1$. It is*

$$h_\gamma(t) = const \cdot e^{-it^{\mathrm{T}}\mu} Q(t)^{-p/2} \phi\left(\tfrac{\gamma-1}{\gamma} Q(t)^{-\frac{\gamma}{2(\gamma-1)}}\right), \quad t \in \mathbb{R}, \tag{25}$$

where $\phi(z)$ is an entire function of z, while

$$const = (\tfrac{\gamma}{\gamma-1})^{p\frac{\gamma-1}{\gamma}} C_\gamma^p \quad \text{and} \quad Q(t) = t^{\mathrm{T}}\Sigma^{-1}t.$$

See [11].

3. Generalized Fisher's information measure for the γ-GND

The Normal distribution was generalized by an extra (shape) parameter $\gamma \in \mathbb{R} \setminus (0,1)$ as introduced in Section 2. The entropy type Fisher's information [19] was also generalized, see [8] with an extra parameter δ. We shall refer hereafter to the δ-GFI. We briefly describe the δ-GFI.

Let X be a multivariate r.v. with p.d.f. f on $\mathbb{R}^{p\times 1}$. For the function $f : \mathbb{R}^{p\times 1} \to \mathbb{R}$ from the Sobolev space $\mathscr{W}^{1,2}(\mathbb{R}^p)$, $\delta > 1$ and for $f^{1/2} \in \mathscr{W}^{1,2}(\mathbb{R}^p)$ Fisher's entropy type information of f, $\mathrm{J}(X)$ is defined by one of the relations below

$$\begin{aligned}\mathrm{J}(X) &= \int_{\mathbb{R}^p} f(x)\|\nabla \log f(x)\|^2 dx = \int_{\mathbb{R}^p} f(x)^{-1}\|\nabla f(x)\|^2 dx \\ &= \int_{\mathbb{R}^p} \nabla f(x)\cdot\nabla \log f(x) dx = 4\int_{\mathbb{R}^p} \|\nabla\sqrt{f(x)}\|^2 dx,\end{aligned}$$

see [8] and [9]. The δ-GFI is defined as

$$\mathrm{J}_\delta(X) = \int_{\mathbb{R}^p} \|\nabla \log f(x)\|^\delta f(x) dx. \tag{26}$$

It is easy then to verify that

$$\mathrm{J}_\delta(X) = \int\limits_{\mathbb{R}^p} \|\nabla f(x)\|^\delta f(x)^{1-\delta} dx = \delta^\delta \int\limits_{\mathbb{R}^p} \|\nabla f(x)^{1/\delta}\|^\delta dx. \tag{27}$$

For the 2-GFI case we are reduced to the usual definition of J, i.e., $\mathrm{J}_2(X) = \mathrm{J}(X)$.

The Blachman-Stam inequality is generalized through δ-GFI. Indeed for given p-variate X and Y independent r.v.'s and $\lambda \in (0,1)$, it holds

$$\mathrm{J}_\delta\left(\lambda^{1/\delta} X + (1-\lambda)^{1/\delta} Y\right) \leq \lambda \mathrm{J}_\delta(X) + (1-\lambda)\mathrm{J}_\delta(Y). \tag{28}$$

The equality holds with X and Y normally distributed with the same covariance matrix, see [8].

Recall that the Shannon entropy H of a r.v. X is defined as, [21],

$$\mathrm{H}(X) = \int\limits_{\mathbb{R}^p} f(x) \log f(x) dx, \tag{29}$$

and therefore the entropy power is defined

$$\mathrm{N}(X) = \nu e^{\frac{2}{p}\mathrm{H}(X)}, \tag{30}$$

with $\nu = (2\pi e)^{-1}$. The extension of the entropy power, the generalized entropy power (δ-GEP) is defined for $\delta \in \mathbb{R} \setminus [0,1]$, as

$$\mathrm{N}_\delta(X) = \nu_\delta e^{\frac{\delta}{p}\mathrm{H}(X)}, \tag{31}$$

where

$$\nu_\delta = \left(\tfrac{\delta-1}{\delta e}\right)^{\delta-1} \pi^{-\delta/2} (\xi_\delta^p)^{\delta/p}, \quad \delta \in \mathbb{R} \setminus [0,1], \tag{32}$$

with ξ_p^δ as in (12). In technical applications, such as signal I/O systems, the generalized entropy power can still be the power of the white Gaussian noise having the same entropy. Trivially, when $\delta = 2$, (31) is reduced to the existing entropy power $\mathrm{N}(X)$, i.e., $\mathrm{N}_2(X) = \mathrm{N}(X)$ as $\nu_2 = \nu$.

Two results are essential due to this extension, see [8].

(i) The information inequality still holds, i.e.,

$$\mathrm{J}_\delta(X)\mathrm{N}_\delta(X) \geq p, \tag{33}$$

(ii) The Cramér-Rao inequality is extended to

$$\left[\tfrac{2\pi e}{p} \operatorname{Var} X\right]^{1/2} \left[\tfrac{\nu_\delta}{p} \mathrm{J}_\delta(X)\right]^{1/\delta} \geq 1. \tag{34}$$

The extended Cramér-Rao inequality (34) under the normality parameter $\delta = 2$, is reduced to the usual Cramér-Rao inequality form

$$\mathrm{J}(X)\,\mathrm{Var}\,X \geq p, \tag{35}$$

see [3]. Moreover, the classical entropy inequality

$$\mathrm{Var}\,X \geq p\mathrm{N}(X) = \tfrac{p}{2\pi e} e^{\frac{2}{p}\mathrm{H}(X)} \quad \text{or} \quad \mathrm{H}(X) \leq \tfrac{p}{2}\log\{\tfrac{2\pi e}{p}\,\mathrm{Var}\,X\}, \tag{36}$$

can be extended into the form

$$\mathrm{Var}\,X \geq p(2\pi e)^{\frac{\delta-4}{\delta}} \nu_\delta^{2/\delta} \mathrm{N}_\delta^{2/\delta}(X) = p(2\pi e)^{\frac{\delta-2}{\delta}} \nu_\delta^{2/\delta} e^{\frac{4}{p\delta}\mathrm{H}_\delta(X)}, \tag{37}$$

through the generalized Shannon entropy H_δ. The H_δ is defined through the generalized entropy power, i.e., $\mathrm{N}_\delta(X) = \nu \exp\{\frac{2}{p}\mathrm{H}_\delta(X)\}$. Under the normality parameter $\delta = 2$, the inequality (37) is reduced to the usual entropy inequality as in (36).

For the γ-GND andδ-GEP we have the following.

Theorem 4. *Let X an elliptically contoured γ-GND r.v. $X \sim \mathcal{N}_\gamma^p(\mu, \Sigma)$. It holds*

$$\mathrm{N}_\delta(X) = \big(\tfrac{\delta-1}{e\delta}\big)^{\delta-1} \big(\tfrac{e\gamma}{\gamma-1}\big)^{\delta\frac{\gamma-1}{\gamma}} \xi_{\delta,\gamma}^p |\Sigma|^{\frac{\delta}{2p}}, \tag{38}$$

where

$$\xi_{\delta,\gamma}^p = \frac{\xi_\delta^p}{\xi_\gamma^p} = \frac{\Gamma\left(p\frac{\gamma-1}{\gamma}+1\right)}{\Gamma\left(p\frac{\delta-1}{\delta}+1\right)}. \tag{39}$$

Corollary 5. *With $\delta = 2$ it holds*

$$\mathrm{N}(X) = \tfrac{1}{2e}\big(\tfrac{e\gamma}{\gamma-1}\big)^{2\frac{\gamma-1}{\gamma}} (\xi_\gamma^p)^{2/p} |\Sigma|^{1/p}, \tag{40}$$

and

$$\mathrm{H}(X) = \log \frac{\pi^{p/2}\xi_\gamma^p\sqrt{|\Sigma|}}{(\delta/e)^{p\delta}}. \tag{41}$$

Theorem 6. *Let X be a spherically contoured γ-GND r.v. $X \sim \mathcal{N}_\gamma^p(\mu, \sigma^2\mathbb{I}_p)$. Then it holds*

$$\mathrm{J}_\delta(X) = p\frac{\Gamma\left(p\frac{p(\gamma-1)+\delta}{\gamma}\right)}{\Gamma\left(p\frac{\gamma-1}{\gamma}+1\right)} \big(\tfrac{\gamma-1}{\gamma}\big)^{\frac{\gamma-\delta}{\gamma}} \sigma^\delta, \tag{42}$$

while the lower and upper bounds for J_δ *are given by*

$$1 < \mathrm{J}_\delta(X) \le \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})}\sqrt{2} < p. \tag{43}$$

See [10] for details.

From (41) we derive that the Shannon entropy for the elliptically countered multivariate Uniform, Normal and Laplace distributed X (i.e., for $\gamma = 1, 2, \pm\infty$ with $X \sim \mathcal{N}^p_\gamma(\mu, \Sigma)$ respectively) is given by

$$\mathrm{H}(X) = \begin{cases} \log \dfrac{\pi^{p/2}\sqrt{|\Sigma|}}{\Gamma\left(\frac{p}{2}+1\right)}, & X \sim \mathcal{N}^p_1(\mu, \Sigma), \\ p \log \sqrt{2\pi e|\Sigma|}, & X \sim \mathcal{N}^p_2(\mu, \Sigma), \\ \log \dfrac{p! e \pi^{p/2}\sqrt{|\Sigma|}}{\Gamma\left(\frac{p}{2}+1\right)}, & X \sim \mathcal{N}^p_{\pm\infty}(\mu, \Sigma), \end{cases} \tag{44}$$

while $\mathrm{H}(X)$ is infinite when $X \sim \mathcal{N}^p_0(\mu, \Sigma) = \mathcal{D}(\mu)$.

Therefore, a global development can be obtained for the information theory approach through the γ-GND development of the Section 2.

As far as the information "distance" is concerned between two γ-GND, [13] worked through the Kullback-Leibler (K-L) measure, [16], of information (also known as relative entropy). Recall that the K-L information $\mathrm{KLI}(f, g)$ between two p-variate density functions f, g is given by

$$\mathrm{KLI}^p_\gamma(f, g) = \int_{\mathbb{R}^p} f(x) \log \frac{f(x)}{g(x)} dx.$$

For appropriate choice of f and g taken from the spherically contoured γ-GND family of distributions, the following results were proven, [13], presented here in the Theorem 7 below.

Let $f_\gamma(x)$ and $g_\gamma(x)$ be the p.d.f. of $\mathcal{N}_\gamma(\mu_1, \sigma_1^2\mathbb{I}_p)$ and $\mathcal{N}_\gamma(\mu_0, \sigma_0^2\mathbb{I}_p)$ respectively. Let us denote with

$$\begin{aligned} q_i(x) &= \tfrac{\gamma-1}{\gamma}\left(\tfrac{1}{\sigma_i}\|x-\mu_i\|\right)^{\frac{\gamma}{\gamma-1}}, \quad x \in \mathbb{R}^p, \quad i = 0, 1 \quad \text{and} \\ E_{i,j}(x) &= \int_{\mathbb{R}^p} e^{-q_i(x)} q_j(x) dx, \quad x \in \mathbb{R}^p, \quad i, j = 0, 1. \end{aligned}$$

Then it holds, see [13],

Theorem 7. *The K-L information measure for the γ-GND defined above, is*

$$\mathrm{KLI}_\gamma^p(f,g) = C_\gamma^p \sigma_1^p \left[p \log \tfrac{\sigma_0}{\sigma_1} \int\limits_{\mathbb{R}^p} e^{q_1(x)} - (E_{1,1} + E_{1,0})(x) dx \right].$$

When f_γ and g_γ are having the same location parameter, i.e., $\mu_0 = \mu_1$, we obtain

$$\mathrm{KLI}_\gamma^p(f,g) = p \log \tfrac{\sigma_0}{\sigma_1} - p(\tfrac{\gamma-1}{\gamma}) \left[1 - (\tfrac{\sigma_1}{\sigma_0})^{\frac{\gamma}{\gamma-1}} \right],$$

while for $\mu_0 \neq \mu_1$ and $\gamma = 2$,

$$\mathrm{KLI}_2^p(f,g) = \tfrac{p}{2} \left[\left(\log \tfrac{\sigma_0^2}{\sigma_1^2} \right) - 1 + \tfrac{\sigma_1^2}{\sigma_0^2} + \tfrac{\|\mu_1 - \mu_0\|^2}{p\sigma_0^2} \right].$$

The $\mathrm{KLI}_2^p(f,g)$ above provides evidence for another interesting extensions through the γ-GND approach.

4. More extensions

We recall that there are cases (for example negative time) where a "truncation" of the Normal distribution is needed. Such cases might be possible either for truncation to the right or to the left. We extend this idea to the γ-GND. Let X be a univariate r.v. from $\mathcal{N}_\gamma(\mu, \sigma^2)$ with p.d.f. f_γ as in (13) and c.d.f. F_γ as in (21). We shall say that X follows the γ-GND truncated to the right at $x = \rho$ with p.d.f. $f_{\gamma;\rho}$ when

$$f_{\gamma;\rho}(x) = \begin{cases} 0, & \text{if} \quad x > \rho, \\ \dfrac{f_\gamma(x)}{F_\gamma(\rho)} = \dfrac{C_\gamma^1(\sigma)}{\Phi_\gamma(\frac{\rho-\mu}{\sigma})} \exp\left\{ -\frac{\gamma-1}{\gamma} \left| \frac{x-\mu}{\sigma} \right|^{\frac{\gamma}{\gamma-1}} \right\}, & \text{if} \quad x \leq \rho, \end{cases} \tag{45}$$

Similarly, it would be truncated to the left at $x = \tau$

$$f_{\gamma;\tau}(x) = \begin{cases} 0, & \text{if} \quad x < \tau, \\ \dfrac{f_\gamma(x)}{1 - F_\gamma(\tau)} = \dfrac{C_\gamma^1(\sigma)}{1 - \Phi_\gamma(\frac{\tau-\mu}{\sigma})} \exp\left\{ -\frac{\gamma-1}{\gamma} \left| \frac{x-\mu}{\sigma} \right|^{\frac{\gamma}{\gamma-1}} \right\}, & \text{if} \quad x \geq \tau, \end{cases} \tag{46}$$

The Lognormal distribution can be also nicely extended to the γ-order Lognormal distribution or γ-GLND, in the sense that if $X \sim \mathcal{N}_\gamma^1(\mu, \sigma^2)$ then e^X will follow the γ-GLND, i.e., $e^X \sim \mathcal{LN}_\gamma(\mu, \sigma)$ with p.d.f.

$$g_\gamma(x) = \tfrac{1}{x} f_\gamma(\log x) = C_\gamma^1(\sigma) x^{-1} \exp\left\{ -\frac{\gamma-1}{\gamma} \left| \frac{\log x - \mu}{\sigma} \right|^{\frac{\gamma}{\gamma-1}} \right\}, \quad x \in \mathbb{R}_+^*. \tag{47}$$

Moreover, if $X \sim \mathcal{LN}_\gamma(\mu, \sigma)$ then $\log X \sim \mathcal{N}_\gamma^1(\mu, \sigma^2)$. Nice results can be also obtained for the γ-GLND.

Interest also might be focused on the quadratic form of γ-GND. Indeed, we state and prove the following.

Theorem 8. *Let $\mathcal{N}_\gamma^p(\mu, \Sigma)$. Then the quadratic form $U = X^{\mathrm{T}}\Sigma^{-1}X$ follows the generalized non-central χ^2 distribution with p degrees of freedom and non-centrality parameter $\lambda^2 = \mu^{\mathrm{T}}\Sigma^{-1}\mu$.*

Proof. The γ-GND family is a subclass of the Kotz type family of distributions $\mathcal{K}(\mu, \Sigma, N, r, s)$ for parameters $N = 1$, $s = \frac{\gamma}{2(\gamma-1)}$ and $r = \frac{\gamma-1}{\gamma}$. Therefore, using a result of [2], $U = X^{\mathrm{T}}\Sigma^{-1}X$ follows the generalized χ^2 as X is a Kotz type distributed r.v. and Theorem has been proved. ■

Notice that, for the trivariate case ($p = 3$), the p.d.f. f_U of $U = X^{\mathrm{T}}\Sigma^{-1}X$, can be expressed as

$$f_U(u) = \frac{(\frac{\gamma-1}{\gamma})^{\frac{\gamma-1}{\gamma}-1}}{2\lambda\,\Gamma(3\frac{\gamma-1}{\gamma})}\left[\Gamma\left(a, 2\tfrac{\gamma-1}{\gamma}\right) - \Gamma\left(b, 2\tfrac{\gamma-1}{\gamma}\right)\right],$$

see [18], where

$$a = (\tfrac{\gamma-1}{\gamma})^{\frac{\gamma}{2(\gamma-1)}}(\sqrt{u}+\lambda)^{\frac{\gamma}{\gamma-1}}, \quad b = (\tfrac{\gamma-1}{\gamma})^{\frac{\gamma}{2(\gamma-1)}}(\sqrt{u}-\lambda)^{\frac{\gamma}{\gamma-1}},$$

while $\Gamma(\cdot,\cdot)$ being the upper incomplete gamma function, [6].

5. Discussion

The Logarithm Sobolev Inequalities (LSI) [22] as well as the Poincaré Inequality (PI) [1] provide food for thought and a solid mathematical framework for Statistics problems, especially when the Normal distribution is involved. Briefly speaking the PI is of the form

$$\mathrm{Var}_\mu(f) \le c_p \int |\nabla f|^2 d\mu, \tag{48}$$

for f differentiable function on $\mathbb{R}^p$ with compact support while μ is an appropriate measure. The constant c_p is known as the Poincaré constant. The Sobolev Inequality is of the form

$$\|s\|_q \le c_s \|\nabla f\|_2, \quad q = \tfrac{2p}{p-2}. \tag{49}$$

The constant c_s is known as the Sobolev constant. Both PI and LSI are applied to Information Theory so that to evaluate bounds on variance, entropy, energy, see [9, 14].

We focused on LSI and working for the appropriate constant c_s, the γ-GND was emerged, discussed in Section 1. The family of the γ-GND was presented in a compact form. There was another attempt to generalize the univariate Normal distribution. The density function of the form

$$h(x) = \frac{b}{2a\,\Gamma(1/b)} \exp\left\{-|\tfrac{x-\mu}{a}|^b\right\}, \quad x \in \mathbb{R}, \tag{50}$$

was the introduced extension, see [17]. This coincides with the γ-GND $\mathcal{N}_\gamma(\mu, \sigma^2)$ for $a = b^b\sigma$ and $b = \gamma/(\gamma-1)$. For the multivariate case see [5]. These existent generalizations is rather technically constructed as it was not obtained through the implementation of a strong mathematical background as the LSI.

Although a number of papers were presented on the generalized Normal we are still investigating more extensions, see and we believe we can cover all the possible applications extending the Normal distribution case.

One of the merits of the family of γ-GND is that includes a number of well known distributions while the singularity of the Dirac distribution being also one of them. Moreover, the extra parameter γ offers, in principle, different shape approaches and therefore heavy-tailed distributions can easily obtained altering parameter γ which effects kurtosis.

In practical problems, such as in Economics where heavy-tailed distributions are needed [4], the γ-GND seems useful. The large positive-ordered GND's provide heavy-tailed distributions as $\mathcal{N}_\gamma^p(\mu, \Sigma)$ approaches the multivariate Laplace distributions, while further heavier-tailed distributions can be extracted through the negative-ordered GND's especially close to zero-ordered GND, i.e., close to the Dirac case. Nevertheless, the higher the dimension gets the heavier the tails become for all multivariate γ-GND's unless we are considering γ-GND's close to the $\mathcal{N}_1^p(\mu, \Sigma)$, i.e., close to the (elliptically contoured) Uniform distribution.

References

[1] W.A. Benter, *Generalized Poincaré inequality for the Gaussian measures*, Amer. Math. Soc. **105** (2) (1989) 49–60.

[2] T. Cacoullos and M. Koutras, *Quadric forms in sherical random variables: Generalized non-central χ^2 distribution*, Naval Research Logistics Quarterly **31** (1984) 447–461. doi:10.1002/nav.3800310310

[3] T.M. Cover and J.A. Thomas, Elements of Information Theory, 2nd Ed. (Wiley, 2006).

[4] K. Fragiadakis and S.G. Meintanis, *Test of fit for asymentric Laplace distributions with applications*, J. of Statistics: Advances in Theory and Applications **1** (1) (2009) 49–63.

[5] E. Gómez, M.A. Gómez-Villegas and J.M. Marin, *A multivariate generalization of the power exponential family of distributions*, Comm. Statist. Theory Methods **27** (3) (1998) 589–600. doi:10.1080/03610929808832115

[6] I.S. Gradshteyn and I.M. Ryzhik, Table of Integrals, Series, and Products (Elsevier, 2007).

[7] L. Gross, *Logarithm Sobolev inequalities*, Amer. J. Math. **97** (761) (1975) 1061–1083. doi:10.2307/2373688

[8] C.P. Kitsos and N.K. Tavoularis, *Logarithmic Sobolev inequalities for information measures*, IEEE Trans. Inform. Theory **55** (6) (2009) 2554–2561. doi:10.1109/TIT.2009.2018179

[9] C.P. Kitsos and N.K. Tavoularis, *New entropy type information measures*, in: Information Technology Interfaces (ITI 2009), Luzar, Jarec and Bekic (Ed(s)), (Dubrovnic, Croatia, 2009) 255–259.

[10] C.P. Kitsos and T.L. Toulias, *Bounds for the generalized entropy-type information measure*, J. Comm. Comp. **9** (1) (2012) 56–64.

[11] C.P. Kitsos, T.L. Toulias and C.P. Trandafir, *On the multivariate γ–ordered normal distribution*, Far East J. of Theoretical Statistics **38** (1) (2012) 49–73.

[12] C.P. Kitsos and T.L. Toulias, *On the family of the γ–ordered normal distributions*, Far East J. of Theoretical Statistics **35** (2) (2011) 95–114.

[13] C.P. Kitsos and T.L. Toulias, *New information measures for the generalized normal distribution*, Information **1** (2010) 13–27. doi:10.3390/info1010013

[14] C.P. Kitsos and T.L. Toulias, *Entropy inequalities for the generalized Gaussia*, in: Information Technology Interfaces (ITI 2010), (Ed(s)), (Cavtat, Croatia, 2010) 551–556.

[15] S. Kotz, *Multivariate distribution at a cross-road*, in: Statistical Distributions in Scientific Work Vol. 1, Patil, Kotz, Ord (Ed(s)), (Dordrecht, The Netherlands: D. Reidel Publ., 1975) 247–270. doi:10.1007/978-94-010-1842-5_20

[16] S. Kullback and A. Leibler A, *On information and sufficiency*, Ann. Math. Statist. **22** (1951) 79-86. doi:10.1214/aoms/1177729694

[17] S. Nadarajah, *A generalized normal distribution*, J. Appl. Stat. **32** (7) (2005) 685–694. doi:10.1080/02664760500079464

[18] S. Nadarajah, *The Kotz type distribution with applications*, Statistics **37** (4) (2003) 341–358. doi:10.1080/0233188031000078060

[19] T. Papaioanou and K. Ferentinos, *Entropic descriptor of a complex behaviour*, Comm. Stat. Theory and Methods **34** (2005) 1461–1470.

[20] M. Del Pino, J. Dolbeault and I. Gentil, *Nonlinear difussions, hypercontractivity and the optimal L^p-Euclidean logarithic Sobolev ineqiality*, J. Math. Anal. Appl. **293** (2) (2004) 375–388. doi:10.1016/j.jmaa.2003.10.009

[21] C.E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J. **27** (1948) 379–423, 623–656. doi:j.1538-7305.1948.tb01338.x

[22] S. Sobolev, *On a theorem of functional analysis*, English translation: AMS Transl. Ser. 2 **34** (1963) 39–68.

Received 20 April 2014
Revised 15 June 2014