Abstract. The target of this paper is to discuss the existent Poincaré and Logarithm Sobolev Inequalities (PI and LSI resp.) for the Gaussian (normal) distribution which is essential in theoretical Statistics and plays an important role in Information Theory and Statistics. The adopted Mathematical background is usually simplified in practical applications. The entropy, energy and variance are related through some order due to PI and LSI. The extended multivariate normal, being a generalized Gaussian, also obeys to LSI.

Keywords. Entropy power, Information measures, Poincaré Inequalities, Logarithmic Sobolev Inequalities.

1. Introduction

The well known normal distribution, introduced by Gauss and therefore also known as Gaussian, plays an important role to all statistical problems. Interest is focused to the Information Theory and Statistics. The Gaussian and the introduced generalized Gaussian (or hyper multivariate normal distribution) has been discussed in [8], while new entropy measures were introduced in [7] and extensively discussed and proved in [6]. Now, the Poincaré and Sobolev Inequalities for the three cases of the Gaussian measure (simple, multivariate, generalized) are presented. Both these classes of inequalities offer a number of bounds for a number of physical applications, the most well known being the Energy, among others.

The Gaussian function or the error function, or the normal distribution, is certainly known, with two parameters – the mean and the variance. For the Gaussian an extra parameter was then introduced in [8], and therefore a generalized form of the Gaussian was obtained. Moreover, the generalized Gaussian is a result of an extremal function for the Sobolev inequality. That is why we believe (Logarithm) Sobolev Inequalities are of interest in this paper. In addition the Poincare inequality offers also the “best” constant for the Gaussian measure, and therefore is of interest to see how Poincare and Sobolev inequalities are acting on the Normal distribution.

In the following section 2 is devoted to recall general results for any function $f$, while section 3 is devoted to the Gaussian. Section 4 is referring to the generalized Gaussian, as in [6] and [7]. Entropy since the time of Clausius, 1865, plays an important role joining physical experimentation and statistical analysis. For the principle of maximum entropy, the normal distribution is essential and eventually is related with the energy and the variance involved. Moreover, the channel capacity it is depending on the entropy since the time of Shannon, 1948. Therefore we would like to know how entropy, energy and variance are related under the Gaussian distribution, for practical problems. To proceed we need a solid mathematical background to cover Statistics and Physics, despite the applicable form of this procedure.

This is why in section 2 these definitions are introduced under a mathematical analysis point of view, while are so applicable (channel capacity etc). Moreover, their relations through inequalities, either Poincaré or Sobolev, are briefly discussed.

Now, let us consider two experiments $E_{X} = (X, \xi)$ and $E_{Y} = (Y, \delta)$ with $X$ and $Y$ being the design spaces and $\xi$, $\delta$ the corresponding design measures from the design spaces $\Xi$ and $\Delta$ respectively, see for details [4]. In practice the design space is where the experimenter performs the experiment and the design measure is,
eventually, due to some mathematical insight, the proportion of the observations devoted for each design point. We shall say that the experiment $E_X$ is sufficient for the experiment $Y$ if there exist a transformation of $X$, say $t(X)$, such that $t(X)$ and $Y$ have identical design measure, or coming from the same distribution. We shall write $E_X \succeq E_Y$. In such a case the Shannon information obtained from $E_X$, say $H_X$, is at least as that obtained in $E_Y$, say $H_Y$, i.e. $H_X \geq H_Y$. Moreover, the same ordering occurs for the Fisher information in terms that $I_X(\theta) - I_Y(\theta)$ is non-negative definite, so $[I_X(\theta)] \succeq [I_Y(\theta)]$, and therefore one could say that D-optimal designs, see [4], as far as for $E_X$ and $E_Y$ concerns, the $E_Y$ is more prefferable.

Now, we consider two experiments, one coming from the Gaussian $N(0, \sigma^2)$ and the other from the Gaussian $N(0, \varphi^2 \sigma^2)$. We say that these experiments are equivalent in terms that the one is sufficient for the other. This is trivially true if we multiply all the observations of the first by $\varphi$, or divide all the observations of the second by $\varphi$.

This is why there is an interest to have at least inequalities among various statistical-analytical measures concerning the Gaussian: to be able to compare the “information” we can obtain for an experiment, which usually is assumed that follows the Gaussian –see sections 3, 4– and is so necessary to the Information Theory.

2. Poincaré and Logarithmic Sobolev Inequalities

The fundamental example of the Gaussian distributions is the starting point to Poincaré Inequalities (PI), see [3], as well as for the Logarithmic Sobolev Inequalities (LSI) as they were defined by the early work of Gross, [5]. Applying these inequalities we obtain the best constants for the PI and LSI inequality for the Gaussian measure on $\mathbb{R}^n$. In what follows we are facing the general case of a probability space. Therefore when we are reduced to more often appeared practical cases (say in $\mathbb{R}$) the theoretical approach still holds, and the appropriate “simplifications” in notation are valid (such as: use “ $dx$ ” in integration).

Let $(\mathbb{E}, F, \mu)$ be a probability space. For each $\mu$-integrable function $f : \mathbb{E} \rightarrow \mathbb{R}$ the expected value $\Exp_{\mu}(f)$ with respect to measure $\mu$ is

$$\Exp_{\mu}(f) := \frac{\int f \, d\mu}{\mu(f)}.$$  

The variance $\Var_{\mu}(f)$ of $\mu$-integrable function $f$ is defined as

$$\Var_{\mu}(f) := \Exp_{\mu}\left(\left[f - \Exp_{\mu}(f)\right]^2\right).$$

The variance is always positive, null if and only if $f$ is $\mu$ -a.e. constant (almost everywhere with respect to $\mu$) and infinite if and only if $f$ is not square $\mu$-integrable.

The entropy $\Ent_{\mu}(f)$ of a $\mu$-integrable positive function $f$ is defined to be

$$\Ent_{\mu}(f) := \Exp_{\mu}(f \log f) - \Exp_{\mu}(f) \log \Exp_{\mu}(f).$$

Applying the inequality $uv \leq u \log u - u + e^v$, $u \in \mathbb{R}_+$, $v \in \mathbb{R}$, the so called variational formula for the entropy is obtained,

$$\Ent_{\mu}(f) := \sup \left\{ \Exp_{\mu}(fg), \; \Exp_{\mu}(e^g) = 1 \right\}. \tag{2}$$

The quantity $\Ent_{\mu}(f)$ is finite if and only if $f \sup(0, \log f)$ is $\mu$-integrable. Notice that when the expected value of $f$ vanishes the definition (1) is simplified. Relation (2) is equivalent to the following inequality, known as entropy inequality.

$$\Exp_{\mu}(fg) \leq \frac{\Exp_{\mu}(f)}{t} \log \Exp_{\mu}(e^g) + \frac{\Ent_{\mu}(f)}{t}, \tag{3}$$

where $f$ is every positive and square integrable function, $g$ is a square integrable function and $t > 0$. The following Proposition 1.1 is referring to the product probability space, as far as its variance and entropy concern. Notice that still we are working with inequalities.

**Proposition 1.1.** Let $(\mathbb{E}, F, \mu, \ i = 1, 2, \ldots, n)$ $n$ probability spaces and $(\mathbb{E}^n, F^n, \mu^n)$ the product probability space. Then,

$$\Var_{\mu^n}(f) \leq \sum_{i=1}^{n} \Exp_{\mu_i}(\Var_{\mu_i}(f)),$$

$$\Ent_{\mu^n}(f) \leq \sum_{i=1}^{n} \Exp_{\mu_i}(\Ent_{\mu_i}(f)).$$

**Proof.** Let a function $g$ defined on $\mathbb{E}^n$ such that $\Exp_{\mu^n}(e^g) = 1$.  

552
there exist a constant
definition of the Poincaré inequality. Indeed:

Both variance and energy are crucial for the
is positive and invariant under the translations.

where
in this paper.

The measure \( \mu \) satisfies the LSI for a certain
function class \( \mathcal{F}_{LS}(\mathbb{E}, \mu) \), if there exists a constant
\( c \in \mathbb{R}^*_+ = (0, +\infty) \) such that

\[
\text{Ent}_{\mu}(f^2) \leq c \text{Ener}_{\mu}(f),
\]

for each function \( f \in \mathcal{F}_{LS}(\mathbb{E}, \mu) \).

Example 2. We may consider \( \mathcal{F}_{LS}(\mathbb{E}, \mu) \) to be
the Sobolev space \( S'(\mathbb{R}^n, \mu) \). The best constant
\( c_{LS}(\mu) \) for the Logarithmic Sobolev Inequality,
for \( f \) not \( \mu \)-a.e. constant, is defined to be

\[
c_{LS}(\mu) = \left\{ \sup \frac{\text{Ent}_{\mu}(f)}{\text{Var}_{\mu}(f)}, \ f \in \mathcal{F}_{LS}(\mathbb{E}, \mu) \right\}^{-1}.
\]

Since,

\[
\text{Ent}_{\mu}(f^2) = \sup \left\{ \text{Exp}_{\mu}(f^2g), \text{Exp}_{\mu}(e^g) = 1 \right\},
\]

we have that the constant is

\[
c_{LS}(\mu) = \sup \left\{ c(g), \text{Exp}_{\mu}(e^g) = 1 \right\},
\]

Where (with \( \text{Ener}_{\mu}(f) > 0 \))

\[
c(g) = \left\{ \sup \frac{\text{Exp}_{\mu}(f^2g)}{\text{Ener}_{\mu}(f)}, \ f \in \mathcal{F}_{LS}(\mathbb{E}, \mu) \right\}^{-1}.
\]

Under some regularity conditions for the measure \( \mu \), the Poincaré inequality as in (5) is

\[
\text{Var}_{\mu}(f) \leq c \int_{\mathbb{E}} |\nabla f|^p \, d\mu,
\]

with \( f \) a differentiable function having compact support (see [1] and the references there). The constant \( c \) is known as Poincaré constant and will be denoted as \( c_p \) in this paper.

In principle, Logarithmic Sobolev Inequalities attempt to estimate the lower order derivatives of a given function in terms of higher order derivatives. The well known Sobolev Inequalities introduced in 1938, see [12] for details. The introductory and well known Logarithmic Sobolev Inequality is

\[
\int_{\mathbb{R}^n} \left| f(x) \right|^{2q} \, dx \leq c \int_{\mathbb{R}^n} \left| \nabla f(x) \right|^2 \, dx,
\]

or, in a compact form, through the norm.

\[
\left\| f \right\|_p \leq c \left\| \nabla f \right\|_q, \quad q = \frac{2}{n}.
\]

The constant \( c \) is known as Sobolev constant and will be denoted as \( c_{LS} \) in this paper. Since then, various attempts were trying to generalize (9). The first optimal Sobolev inequality was of the form (with \( p \in [1, n] \))

\[
\int_{\mathbb{R}^n} \left| f(x) \right|^{2p} \, dx \leq C_{n, p} \left( \int_{\mathbb{R}^n} \left| \nabla f(x) \right|^p \, dx \right)^{\frac{2}{p}}.
\]
Recall the inequalities (5), (7), (9) and (10). These inequalities are depending on a constant \( c \), which we try to evaluate, on the optimal sense as in (6) and (8) for the PI and LSI respectively.

Therefore, in all these integral inequalities the crucial points are: the exponent, and the value of the critical constant \( c \), which is usually depending on the gamma function. This is clear on the generalized form of Gaussian, introduced in [8] and discussed in [6] and [7].

### 3. PI and LSI for the Gaussian

One of the merits that Gaussian distribution offers to the Information Theory is that for any random variable \( X \) and the estimator \( \text{est}(X) \) the following inequality holds.

\[
\text{Exp}(X - \text{est}(X))^2 \geq (2\pi e)^{-1} \exp\{2h(X)\},
\]

with \( h(X) \) being the differential entropy, see [2]. The equality holds if and only if \( X \) is Gaussian and \( \text{est}(X) \) is the mean of \( X \). This very useful result can be also extended even when side information is given for the estimator, see [2].

Moreover, the Gaussian distribution is adopted for the noise acting additively to the input variable when an input-output time discrete channel is formed. Therefore, the Gaussian distribution needs a special treatment evaluating Poincaré and Sobolev Inequalities.

Both the Poincaré and Logarithmic Sobolev Inequalities are applied to Statistical distributions so that to evaluate bounds between variance, entropy and energy. Moreover, the development of the Poincaré Inequality and Logarithmic Sobolev Inequality for the Gaussian depends on the development on the Bernoulli measure due to a theoretical insight, which is not presented here. Therefore, we discuss firstly the Bernoulli case.

If \( \mathbb{E} = \{0,1\} \) the Bernoulli measure \( \beta_p \) of \( \mathbb{E} \) with the parameter \( p \in (0,1) \) is the following probability measure

\[
\beta_p := p\delta_0 + q\delta_1,
\]

where \( q = 1 - p \) and \( \delta_0 \) is the Dirac measure at \( 0 \).

We evaluate \( \text{Exp}_{\beta_p}(f) = pf(0) + qf(1) \) and the energy is evaluated to be \( \text{Ener}_{\beta_p}(f) = pq|f(0) - f(1)|^2 \). A simple calculation gives \( \text{Var}_{\beta_p}(f) = \text{Ener}_{\beta_p}(f) \) that leads to the Poincaré Inequality for the Bernoulli measure:

**Theorem 3.1.** (PI for Bernoulli measure)

\[
\text{Var}_{\beta_p}(f) \leq \text{Ener}_{\beta_p}(f), \text{ i.e. } c_p(\beta_p) = 1.
\]

**Proof.** See [9].

Next we give the sharp Logarithmic Sobolev Inequality for Bernoulli measure, so that to be clear the application and the comparison between the continuous and the discrete case.

**Theorem 3.2.** (LSI for Bernoulli measure) The best constant for the inequality

\[
\text{Ent}_{\beta_p}(f^2) \leq c_{LS} \text{Ener}_{\beta_p}(f), \quad \text{(13)}
\]

\[
\text{c}_{LS} = \begin{cases} 2, & \text{if } p = \frac{1}{2} \\ \log - \log p, & \text{otherwise.} \end{cases}
\]

Notice that the constant \( c_{LS} \) is a concave function of the parameter \( p \). It diverges to \( +\infty \) as \( p \) tends to 0 and has minimum for \( p = 1/2 \), (as one could expect for the Bernoulli trials) and then the constant depends only on the parameter \( p \). Therefore, considering \( \mathbb{E} = \{a, b\} \) and \( \beta_p := p\delta_a + q\delta_b \) we have the same constant for the inequality. In this case the energy is evaluated.

\[
\text{Ener}_{\beta_p}(f) = pq|f(b) - f(a)|^2.
\]

**Proof.** See [9].

Using the tensorisation property of variance and entropy we obtain the PI as well as the LSI for Gaussian measure from the above inequalities for the Bernoulli measure. Let \( \mathbb{E} = \mathbb{R} \). The Gaussian probability measure is

\[
d\gamma = (2\pi)^{-\frac{n}{2}}e^{\frac{1}{2}d^2\gamma} dx.
\]

**Theorem 3.3.** (PI for the Gaussian on \( \mathbb{R} \)) For \( f \in S_1(\mathbb{R}, \gamma) \):

\[
\text{Var}(f^2) \leq \text{Ener}(f), \quad \text{i.e. } c_\gamma(\gamma) = 1,
\]

**Proof.** See [9].

**Theorem 3.4.** (LSI for the Gaussian on \( \mathbb{R} \)) For \( f \in S_1(\mathbb{R}, \gamma) \):

\[
\text{Ent}(f^2) \leq 2\text{Ener}(f), \quad \text{i.e. } c_{LS}(\gamma) = 2.
\]

**Proof.** The proof is a step by step transfer of the proof of Theorem 3.3 using the tensorisation property of entropy.

Let \( \mathbb{E} = \mathbb{R}^n \) and the Gaussian probability measure on \( \mathbb{R}^n \), \( d\gamma^{\otimes n}(x) = (2\pi)^{-\frac{n}{2}}e^{\frac{1}{2}d^2\gamma} dx \).

The next Theorem 3.5 gives the best constants for the Poincaré and Logarithmic Sobolev Inequality for the Gaussian measure on \( \mathbb{R}^n \), i.e. for the variance of \( f \) and the entropy of \( f^2 \).

Using the following result
Ener_{\nu,\sigma}(f) = \exp_{\nu,\sigma}(\|\nabla f\|^2) = \sum_{i=1}^{n} \exp_{\nu,\sigma}(\|\varphi_i(f, t)\|) = 
\sum_{i=1}^{n} \exp_{\nu,\sigma}\left(\exp_{\nu}(\|\varphi_i(f)\|)\right) ,
\sum_{i=1}^{n} \exp_{\nu,\sigma}\left(\exp_{\nu}(\|\varphi_i(f)\|)\right),
we can deduce from 3.4 the Poincaré and Logarithmic Sobolev Inequality for Gaussian measure on \( \mathbb{R}^n \). It is interesting to notice the simplicity of the involved constants, with values 1 and 2, for PI and LSI respectively. Then:

**Theorem 3.5.** (PI and LSI for Gaussian measure on \( \mathbb{R}^n \)) For \( f \in \mathcal{S}(\mathbb{R}^n, r_{\sigma,\nu}) \) the following are true:

\[ \text{Var}_{\nu,\sigma}(f) \leq \text{Ener}_{\nu,\sigma}(f) \text{, i.e. } c_{\nu,\sigma}(r_{\sigma,\nu}) = 1 \]

\[ \text{End}_{\nu,\sigma}(f^2) \leq 2\text{Ener}_{\nu,\sigma}(f) \text{, i.e. } c_{\nu,\sigma}(r_{\sigma,\nu}) = 2 . \]

Notice that the values of the constants, as it has already mentioned, are rather nice and easy to be adopted in applications, as the involved constants for the multivariate normal discussed below, see relations (20), (21). Therefore there is a simplification in the real life problems.

Now consider the multivariate normal distribution, with mean \( \mu \) and covariance matrix \( \Sigma \) of the form:

\[ N(\mu, \Sigma) = (2\pi)^{-n/2} |\det \Sigma|^{-1/2} \times \exp\left(-\frac{1}{2}(x - \mu)^{\top} \Sigma^{-1}(x - \mu)\right) , \]

with \( \langle a, b \rangle \) being the inner product of the vectors \( a, b \). In this general case of the Gaussian measure the Poincaré and Logarithmic Sobolev Inequality are the following:

\[ \text{Var}_{\nu,\sigma}(f) \leq \sigma^2 \text{Exp}_{\nu,\sigma}(\|\nabla f\|^2) , \quad (20) \]

\[ \text{Ent}_{\nu,\sigma}(f^2) \leq 2\sigma^2 \text{Exp}_{\nu,\sigma}(\|\nabla f\|^2) , \quad (21) \]

respectively.

Now, as far as the entropy of a random vector \( X \) concerns, \( H(X) \) say, considering the following proposition a bound for it is obtained, depending only on the covariance matrix.

**Proposition 3.1.** Let the random vector \( X \) has zero mean and covariance \( \Sigma \). Then

\[ H(X) \leq \frac{1}{2} \log \left(2\pi e\right) |\det \Sigma| \]

with equality if and only if \( X \sim N(0, \Sigma) \).

Proof. See [2].

This Proposition is crucial and clarifies that the entropy for the Gaussian is depending, eventually, only on the variance-covariance matrix, while equality holds when \( X \) is following the (multivariate) normal distribution, a result quite often applied in engineering problems, and information systems.

4. LSI and the Generalized Gaussian

As a result from relation (20), for the particular case of the Gaussian measure, the LSI can be reformed in terms of the following way of thinking: For the variance \( \sigma^2 > 0 \) set \( \sigma' = \sigma \) and let the Gaussian law to be with the covariance matrix \( \sigma^2 I_n = \text{diag}(\sigma, \sigma',..., \sigma') \). In such a case

\[ \text{Ent}_{\nu,\sigma}(f^2) \leq 2\sigma^2 \text{Exp}_{\nu,\sigma}(\|\nabla f\|^2) . \]

With the special form of the function \( f(x) = \exp\left(\frac{1}{2\sigma^2} \|x\|^2\right) \) (which can be considered and Gaussian) the above is reduced to

\[ \int_{\mathbb{R}^n} \log |g(x)|^2 \, dx \leq \frac{1}{2} \log \left(\frac{\sigma}{\sigma'}\right) \int_{\mathbb{R}^n} |g(x)|^2 \, dx . \]

Consider the multiple of normal distribution given from \( (\Sigma = \frac{\sigma^2}{\sigma^2} I_n, \Sigma^{-1} = \frac{\sigma^2}{\sigma^2} I_n) \)

\[ (2\pi)^{-n/4} N(\mu, \frac{\sigma^2}{\sigma^2} I_n) = (2\pi)^{-n/4} (2\pi)^{n^2} (\frac{\sigma^2}{\sigma^2})^{-n^2/2} \times \exp \left(-\frac{1}{2\sigma^2} \|x - \mu\|^2\right) . \]

Then we can state a nice property of the Gaussian distribution, as far as the LSI concerns.

**Theorem 4.1.** The multiple of the Gaussian as in (23) corresponds to the extremals functions for (22).

Proof. See [6].

We define the following generalized Gaussian distribution with mean \( \mu \), covariance matrix \( \Sigma \) and density function of the form

\[ K_{\alpha}(m, \Sigma) = C(n, \alpha) |\det \Sigma|^{-1/2} \times \]

\[ \exp \left(-\frac{1}{2} <(X - m)^\top \Sigma^{-1}(X - m)\right) . \]

The normality factor \( C(n, \alpha) \) is defined to be:

\[ C(n, \alpha) = \pi^{-n/2} \frac{\Gamma \left(\frac{n}{2} + 1\right)}{\Gamma \left(\frac{n}{2} + \frac{1}{\alpha} + 1\right)} . \]

See [6] and [8] for details. Obviously,

\[ K_{\alpha}(m, \Sigma) = 2^{n/2} \pi^{-n/2} |\det \Sigma|^{-1/2} N(m, \Sigma)^{\alpha} . \]

It can be proved that \( K_{\alpha}(m, \Sigma) \) is related to the Kotz type distribution, see [6] for details. Moreover, the function

\[ K_{\alpha}(m, \frac{\sigma'\Sigma}{\sigma}\|I_n\|^{\alpha}) \]

corresponds to extremals function for inequality extending LSI due to Del Pino and Dolbeaut [11].
For the proof see [6]. The essential result is that the defined generalized (hyper) multivariate normal distribution works as an extremal function to a generalized form of Sobolev Inequality.

5. Comments

In this paper we discuss for the $\mu$-integrable function $f$, the expected value $\text{Exp}_\mu(f)$, the entropy $\text{Ent}_\mu(f)$ and the energy $\text{Ener}_\mu(f)$ and their relations through the Poincaré Inequality and the Logarithm Sobolev Inequality. Although the introduced framework due to PI and LSI is generally referred to “constants”, the optimal evaluation of these constants results rather easy and applicable constants, at least for the Gaussian distribution.

Due to this analysis the variance is bounded for the energy see (5), which also bounds the entropy of $f^2$, see (7). When the measure $\mu$ is reduced to Bernoulli relations (5) and (7) still holds, under some assumptions, see (12) and (13) with the constant taking the appropriate value under the Poincaré or the Sobolev framework. For the Gaussian probability measure $d\gamma$, see (14), which appears an aesthetic appeal in practical applications, the same inequalities are transferred as in (15), (16) and (17), (18), while an extension occurs at (20) and (21).

Then, the multivariate normal distribution with the form (19) corresponds to the extremal function due to LSI theory as it was proved in Proposition 5.1 and therefore a generalized Gaussian as in (24), being member of the Kotz type family of distributions, [10], is acting as an extremal for the LSI.

Our target was not to present the proofs but to collect and present general results on inequalities. A great number of other, more applicable perhaps, inequalities concerning the Entropy and the Information Theory, are in [2] ch. 17, were the role of inequalities is highly recognized. We tried to move under this line of thought, so useful to applications on information systems, extending the existed ideas.

This paper attempts to improve the idea of adopting inequalities concerning Expectation, Entropy, Energy and Information for the Statistical Information approach, offering certain bounds, so that to be clear, adopting a theoretical framework, what the optimal value can be in practice.

Acknowledgements

We would like to thank the referee for the useful comments, which help us to improve it.

References