On the Extremal Case of Gaussian Inequality

Syeda Rabab Mudakkar

Center for Mathematical and Statistical Sciences
Lahore School of Economics, Pakistan
drrabab@lahoreschool.edu.pk

Abstract

This work is motivated by comparison and characterization problems for sums of symmetric independent random variables in probability theory. We characterize the particular function, such that the functional of a sum of mixture of independent, identically distributed (i.i.d) Gaussian random variables can be represented via functional of sum of non-mixture i.i.d Gaussian random variables.

Mathematics Subject Classification: 60E15, 60F20

Keywords: Gaussian random variable, characterization, extremal problem

1 Introduction and Preliminaries

One of the key question in comparison problems of probability theory is to describe the distributions of sums of independent random variables via a few simple functionals of their terms. In case of limit theorems, rate of convergence etc, the theory is well established (see, for example, Petrov [7]). However, about the optimal explicit inequalities, not much is known still. This work is motivated by [4, 3, 5, 6, 2] etc, who obtained the sharp upper and lower bounds for sums of symmetric independent random variables. More exactly, we are interested to consider the extremal case of a sum of mixtures of independent, identically distributed Gaussian random variables i.e. we want to characterize \( g(X) \), for which the following identity holds:

\[
ET\left( \sum_{i=1}^{n} Z_i g(X_i) \right) = ET\left( \sum_{i=1}^{n} Z_i g(EX_i) \right)
\]  

(1.1)

where \( \{X_i\}_{i=1}^{n}, \{Z_i\}_{i=1}^{n} \) be two mutually independent sequences of random variables in a space \( \mathbb{R} \), \( (X_i)^n_{i=1} \) is any independent sequence of non-negative
random variables i.e. \( P(X_i \geq 0) = 1 \) and \((Z_i)_{i=1}^n\) is an independent, identically distributed sequence of Gaussian random variables. Moreover, \( f(x) \) belong to a fractional power function or an exponential family.

Our approach of construction is adopted from Whittle [4] as developed in Utev [5]. The result uses an operator approach and in this regard \( A_X f(t) \) denotes the convolution operator applied to the class of fractional power function and an exponential function at point \( t \). The usual second derivative test of the function \( \pi_t(y) = \pi_{g,T}(y;t) = T(t + g(y)) \) characterise \( g_t(y) \) for which \( \pi_t(y) \) is linear on positive real line. If \( \pi_t(y) \) is linear on \([0, \infty)\), for certain \( t \in \mathbb{R}^+ \), then for every non-negative random variable \( X \) with \( E(X) < \infty \), using Jensen’s inequality, we get \( E(\pi(X)) = \pi(EX) \).

We also need the following property of 2-stable (normal) random variable to prove our results:

**Property 1.1 (Nolan [1]).** Let \((Z_i)_{i=1}^n\) is an independent, identically distributed sequence of Gaussian random variables and \((X_i)_{i=1}^n\) is any sequence of non-negative random variables, then

\[
\sum_{i=1}^n Z_i X_i =^d Z \|X\|_2,
\]

where \( \|X\|_2 = \left( \sum_{i=1}^n X_i^2 \right)^{\frac{1}{2}} \).

## 2 Main Results

We start with the following simple Lemma, which provides a general framework.

**Lemma 2.1.** Fix any measurable function \( T \) and \( g \) and define \( \pi_t(y) = \pi_{g,T}(y;t) = ET(t + Z g(y)) \). If \( \pi_t(y) \) is linear for all \( t, y \), then the identity (1.1) holds.

**Proof.** Let

\[
S_k = \sum_{i=1}^k Z_i g(X_i) , U_k = \sum_{i=k}^n Z_i g(EX_i) ,
\]

\[
Q_k = S_k + Z_{k+1} g(X_{k+1}) + U_{k+2} ,
\]

where \( U_k \) and \( g(X_k) \) equal zero for \( k > n \).

Assume linearity of \( \pi_t(y) \) and observe that by conditioning on \( S_k + U_{k+2} \),

\[
ET(Q_k) = ET(S_k + Z_{k+1} g(X_{k+1}) + U_{k+2})
\]

\[
= ET(S_k + Z_{k+1} g(EX_{k+1}) + U_{k+2})
\]

\[
= ET(Q_{k-1}) ,
\]
On the extremal case of Gaussian inequality

and so by induction

\[ ET\left(\sum_{i=1}^{n} Z_i g(X_i)\right) = ET(Q_n) \]

\[ = ET(Q_0) = ET\left(\sum_{i=1}^{n} Z_i g(EX)\right). \]

Next we present our result for the class of fractional power function.

**Theorem 2.2.** Let \((a_i)_{i=1}^{n}\) is a sequence of real numbers and \(\|a\|_2\) represents the \(L_2\) norm. Assume \(\phi_t(y) = \left(t + g^2(y)\right)^{\frac{p}{2}}\) is linear on positive real line. Then for any positive \(p\), the identity

\[ E\left|\sum_{i=1}^{n} Z_i g(X_i)\right|^p = E|Z|^p \|g(a)\|_2^p \]  

(2.1)

holds.

**Proof.** Clearly the Property 1.1 implies that

\[ E\left[|\sum_{i=1}^{n} Z_i g(X_i)|^p\right] = E\left[|Z|^p \sum_{i=1}^{n} g^2(X_i)\right]^{\frac{p}{2}}. \]

Next consider a convolution operator

\[ A_X f(t) = Ef(t + Zg(X)), \]

where \(f(x) = x^\frac{p}{2}\) which implies

\[ A_X f(t) = E\left[|Z|^p [t + g^2(X)]^\frac{p}{2}\right]. \]

Consider

\[ \pi_t(y) = \left(E|Z|^p\right)\left(t + g^2(y)\right)^{\frac{p}{2}}. \]

So without loss of generality, we work with the function

\[ \phi_t(y) = \left(t + g^2(y)\right)^{\frac{p}{2}}. \]

Assuming linearity of \(\phi_t(y)\) and using Jensen’s inequality for a non-negative random variable \(X\), we deduce that

\[ E(\phi_t(X)) = \left(t + g^2(X)\right)^{\frac{p}{2}} = \left(t + g^2(EX)\right)^{\frac{p}{2}} = \phi_t(EX) \quad \text{for every} \quad t \in \mathbb{R}^+. \]
Consider argument of Lemma 2.1 i.e. conditioning on all \( X_i \)’s except one and assuming that \( X_i \)’s are independent, we obtain
\[
\prod_{i=1}^{n} A_{X_i}f(t) = E\left[ |Z|^p \left( t + \sum_{i=1}^{n} g^2(X_i) \right)^{\frac{p}{2}} \right] \\
= E\left[ |Z|^p \left( t + \sum_{i=1}^{n} g^2(EX_i) \right)^{\frac{p}{2}} \right] = \prod_{i=1}^{n} A_{EX_i}f(t).
\]

Thus
\[
\prod_{i=1}^{n} A_{X_i}f(0) = E\left[ |Z|^p \left( \sum_{i=1}^{n} g^2(X_i) \right)^{\frac{p}{2}} \right] \\
= E\left[ |Z|^p \left( \sum_{i=1}^{n} g^2(EX_i) \right)^{\frac{p}{2}} \right] = \prod_{i=1}^{n} A_{EX_i}f(0).
\]

Letting \( a_i = E(X_i) \) implies the required result.

Next we present our result that characterize the linearity of \( \phi_t(y) \).

**Theorem 2.3.** The function \( \phi_t(y) = \left( t + g^2(y) \right)^{\frac{p}{2}} \) is linear on positive real line, if and only if, for any fixed positive \( p \)
\[
g_t(y) = \left[ \left( pe^C y + \frac{p}{2} D \right)^{\frac{2}{p}} - t \right]^{\frac{1}{2}}
\]
where \( C \) and \( D \) are arbitrary constants of integration.

**Proof.** Clearly to prove the necessary condition, we need to consider \( \phi''_t(y) = 0 \). We omit the details, because the arguments involve extensive calculus including integration by parts, method of partial fractions and substitution etc. We left the proof of sufficient condition because it is obvious from identity (2.2).

**Remark 2.4.** The last two results implies that the identity (2.1) holds, when
\[
g(X) = \left( pe^C X + \frac{p}{2} D \right)^{\frac{1}{p}}
\]
Next, we characterize the value of \( p \) for which \( g_t(y) \) can exists.

**Lemma 2.5.** Assume the domain of \( g_t(y) \) consider three different points of \( y \) i.e. \( y_0, y_1 \) and \( y_2 \). Let \( g_t^2(y_0) = A, g_t^2(y_1) = B \) and \( g_t^2(y_2) = Q \), where \( A, B \) and \( Q \) are three distinct real constants. Then \( g_t(y) \) exists only when \( p = 2 \), for all positive values of \( t \).
Proof. Letting $pe^C = C_1$ and $\frac{pD}{2} = C_2$, the set of equations obtained from eq. (2.2) can be rewritten as

\begin{align}
(A + t)^{\frac{p}{2}} &= C_1y_0 + C_2, \\
(B + t)^{\frac{p}{2}} &= C_1y_1 + C_2, \\
(Q + t)^{\frac{p}{2}} &= C_1y_2 + C_2.
\end{align}

Subtracting eq. (2.4) from eq. (2.3), we obtain

$$C_1 = \frac{(A + t)^{\frac{p}{2}} - (B + t)^{\frac{p}{2}}}{y_0 - y_1}.$$

Similarly subtracting eq. (2.4) from (2.5), we obtain

$$C_1 = \frac{(Q + t)^{\frac{p}{2}} - (B + t)^{\frac{p}{2}}}{y_2 - y_1},$$

hence

$$\frac{(A + t)^{\frac{p}{2}} - (B + t)^{\frac{p}{2}}}{y_0 - y_1} = \frac{(Q + t)^{\frac{p}{2}} - (B + t)^{\frac{p}{2}}}{y_2 - y_1} \quad (2.6)$$

Let $d = \frac{1}{y_0 - y_1}$, $e = \frac{y_0 - y_2}{(y_0 - y_1)(y_2 - y_1)}$ and $f = \frac{1}{y_1 - y_2}$, thus

$$d(A + t)^{\frac{p}{2}} + e(B + t)^{\frac{p}{2}} + f(Q + t)^{\frac{p}{2}} = 0 \quad \text{for all} \quad t \geq 0,$$

which is only possible when $p = 2$. \hfill \square

Remark 2.6. Assume the domain of function $g_t(y)$ consider $n$ different points of $y$ i.e. $y_0, y_1, y_2, \ldots, y_{n-1}$ and the image of the square of function consider $n$ distinct points i.e. $A_0, A_1, \ldots, A_{n-1}$. Then $g_t(y)$ exists only when $p = 2$, for all positive $t$.

The following result explains that two points in the image of $g_t(y)$ is not enough to guarantee the characterization. We prove it by the counter example for a particular case of two points i.e. $0$ and any other real point.

Lemma 2.7. Assuming $g_t^2(y) = A$ for $y \in \mathbb{R}$ and $g_t^2(y) = B$ for $y \in \mathbb{R}^c$, and $A \neq B$. Then $g_t^2(y)$ does not exists for any $t \geq 0$.

Proof. On the contrary, assume

$$g_t^2(y) + t = \left[p \left(e^C y + \frac{D}{2}\right)\right]^{\frac{p}{2}}.$$
Let \( R = \{0\} \), which implies (refering to Lemma 2.5), \( y_0 = 0 \) and \( y_1 \neq y_2 \neq 0 \), where \( g^2(0) = A \), \( g^2(y_1) = B \) and \( g^2(y_2) = Q = B \).

Recall equation (2.6), thus

\[
\frac{(A + t)\hat{p}_2 - (B + t)\hat{p}_2}{y_1 - y_2} = \frac{(B + t)\hat{p}_2 - (B + t)\hat{p}_2}{y_2 - y_1},
\]

hence

\[
(A + t)\hat{p}_2 = (B + t)\hat{p}_2,
\]

which is not possible because \( A \neq B \).

Next we present our result when \( T \) belongs to the class of exponential function.

**Theorem 2.8.** Let \((Z_i)_{i=1}^n\) is a sequence of independent, identically distributed standard normal random variables. Assume \( \pi_t(y) := E[e^{\alpha(t + zg(y))}] \) is linear. Then for any fixed \( \alpha \in \mathbb{R} \), the identity

\[
E\left[ \exp\left( \alpha \left( \sum_{i=1}^n Z_i g(X_i) \right) \right) \right] = E\left[ \exp\left( \alpha \left( \sum_{i=1}^n Z_i g(EX_i) \right) \right) \right]
\]

holds.

**Proof.** The proof follows the same lines as Theorem 2.2, except we need to consider a convolution operator \( A_X f(t) = E[e^{\alpha(t + Zg(X))}] \).

**Theorem 2.9.** The function \( \Phi(y) = E[e^{\alpha(zg(y))}] \) is linear on positive real line, when

\[
g(y) = [A + \ln(y)]^{\frac{1}{\alpha}},
\]

where \( A \) is the constant of integration.

**Proof.** We want to characterize linearity of the following function

\[
\pi_t(y) := E[e^{\alpha(t + zg(y))}] = e^{\alpha t} E[e^{\alpha zg(y)}].
\]

So without loss of generality, we can work with the function

\[
\Phi(y) := E[e^{\alpha zg(y)}].
\]

Similar to Proof of Theorem 2.3, we consider

\[
\Phi''(y) = \left( E(z^2 e^{\alpha zg(y)}) (g'(y))^2 + E(z e^{\alpha zg(y)}) g''(y) \right).
\]

Now

\[
E(z e^{\alpha zg(y)}) = \frac{1}{\alpha} E\left( e^{\alpha zg(y)} \right)'_{g(y)} \quad \text{and} \quad E(z^2 e^{\alpha zg(y)}) = \frac{1}{\alpha^2} E\left( e^{\alpha zg(y)} \right)''_{g(y)},
\]
hence
\[ \Phi''(y) = \left[ \left( E(e^{\alpha z g(y)}) \right)''_{g(y)} (g'(y))^2 + \left( E(e^{\alpha z g(y)}) \right)'_{g(y)} \right]. \]

Next
\[ Z \sim N(0,1), \] which implies \[ E(e^{\alpha z g(y)}) = e^{\frac{\alpha^2 (g(y))^2}{2}}. \]

Then
\[ \Phi''(y) = e^{\frac{\alpha^2 (g(y))^2}{2}} \left[ \alpha^2 (g'(y))^2 + \alpha^4 (g'(y))^2 (g(y))^2 + \alpha^2 g(y) g''(y) \right]. \]

Consider \( \Phi''(y) = 0 \), clearly \( e^{\frac{\alpha^2 (g(y))^2}{2}} \neq 0 \), hence
\[ \left[ \alpha^2 (g'(y))^2 + \alpha^4 (g'(y))^2 (g(y))^2 + \alpha^2 g(y) g''(y) \right] = 0. \]

Using integration by parts and substitution, we obtain
\[ g(y) = [A + \ln(y)]^{\frac{1}{\alpha^2}} \text{ where } A = \ln(B). \]

\[ \square \]

**Remark 2.10.** The identity (2.7) holds when \( g(X) = [A + \ln(X)]^{\frac{1}{\alpha^2}}. \)

### 3 Conclusions

We characterized the particular function, that represented the functional of a sum of mixtures of i.i.d Gaussian random variables via functional of sum of non-mixture i.i.d Gaussian random variables. We achieved this goal for fractional power class and an exponential class of functions. In case of power class, we also indicated the situations when the solution of characteristic function exist. However, the idea worked only with these two classes of functions. Thus, the problem for sums of i.i.d Gaussian random variables is solved.

### Acknowledgement

The author would like to thanks Dr. Sergey Utev for his invaluable guidance and support. Thanks are also due to Higher Education Commission of Pakistan and Lahore School of Economics for funding this research.
References


Received: March, 2012