



Complete monotonicity of the entropy in the central limit theorem for gamma and inverse Gaussian distributions

Yaming Yu*

Department of Statistics, University of California, Irvine 92697, USA

ARTICLE INFO

Article history:

Received 16 June 2008

Received in revised form 11 August 2008

Accepted 14 August 2008

Available online 22 August 2008

ABSTRACT

Let $H_g(\alpha)$ be the differential entropy of the gamma distribution $\text{Gam}(\alpha, \sqrt{\alpha})$. It is shown that $(1/2) \log(2\pi e) - H_g(\alpha)$ is a completely monotone function of α . This refines the monotonicity of the entropy in the central limit theorem for gamma random variables. A similar result holds for the inverse Gaussian family. How generally this complete monotonicity holds is left as an open problem.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The classical central limit theorem (CLT) states that for a sequence of independent and identically distributed (i.i.d.) random variables X_i , $i = 1, 2, \dots$, with a finite mean $\mu = E(X_1)$ and a finite variance $\sigma^2 = \text{Var}(X_1)$, the normalized partial sum $Z_n = (\sum_{i=1}^n X_i - n\mu)/\sqrt{n\sigma^2}$ tends to $N(0, 1)$ in distribution as $n \rightarrow \infty$. There exists an information theoretic version of the CLT. The differential entropy for a continuous random variable S with density $f(s)$ is defined as

$$H(S) = - \int_{-\infty}^{\infty} f(s) \log f(s) ds.$$

Barron (1986) proved that if the (differential) entropy of Z_n is eventually finite, then it tends to $(1/2) \log(2\pi e)$, the entropy of $N(0, 1)$, as $n \rightarrow \infty$. As a consequence of Pinsker's inequality (Cover and Thomas, 1991), convergence in entropy implies convergence in distribution in the normal CLT case.

An interesting feature of the information theoretic CLT is that $H(Z_n)$ increases monotonically in n (which reminds us of the second law of thermodynamics); since the entropy is invariant under a location shift, equivalently for a sequence of i.i.d. random variables X_i with a finite second moment, we have

$$H\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i\right) \leq H\left(\frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} X_i\right), \quad n \geq 1. \quad (1)$$

Although the special case $n = 1$ is known to follow from Shannon's classical entropy power inequality (Stam, 1959; Blachman, 1965), it was not until 2004 before the above inequality was finally proven by Artstein et al. (2004); see also Madiman and Barron (2007) for generalizations.

This paper is concerned with a possible refinement of this monotonicity, at least when the distribution of X_i belongs to some special families. Recall that a function $f(x)$ on $x \in (0, \infty)$ is called *completely monotonic* if its derivatives of all orders exist and satisfy

$$(-1)^k f^{(k)}(x) \geq 0, \quad x > 0,$$

* Tel.: +1 949 824 7361.

E-mail address: yamingy@uci.edu.

for all $k \geq 0$. In particular, a completely monotone function is both monotonically decreasing and convex. A discrete sequence f_n on $n = 1, 2, \dots$ is completely monotonic if

$$(-1)^k \Delta^k f_n \geq 0, \quad n = 1, 2, \dots,$$

for all $k \geq 0$, where Δ denotes the first difference operator, i.e., $\Delta f_n = f_{n+1} - f_n$. Some basic properties of completely monotone functions can be found in Feller (1966).

Theorem 1. Let X_i , $i = 1, 2, \dots$, be i.i.d. random variables with distribution F , mean μ , and variance $\sigma^2 \in (0, \infty)$. Denote

$$h(n) = \frac{1}{2} \log(2\pi e) - H \left(\frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^n (X_i - \mu) \right), \quad n = 1, 2, \dots$$

Then $h(n)$ is a completely monotonic function of n if F is either a gamma distribution or an inverse Gaussian distribution.

Note that in Theorem 1, $h(n)$ is the also the relative entropy, or Kullback–Leibler divergence, between $(1/\sqrt{n\sigma^2}) \sum_{i=1}^n (X_i - \mu)$ and the standard normal. Theorem 1 indicates that the behavior of relative entropy in the classical CLT is (at least for gamma and inverse Gaussian families) highly regular in a sense.

In Sections 2 and 3, we derive Theorem 1 for the gamma and inverse Gaussian families respectively. Part of the reason these two families are chosen is that they are analytically tractable. It seems intuitive that Theorem 1 should hold for a wide class of distributions, but the proofs presented here depend heavily on properties of the gamma and inverse Gaussian distributions, and are unlikely to work for the general situation. How Theorem 1 can be extended is an open problem.

2. The gamma case

Let $H_g(\alpha, \beta)$, $\alpha > 0$, $\beta > 0$, be the entropy of a $\text{Gam}(\alpha, \beta)$ random variable whose density is specified by

$$p(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x}, \quad x > 0,$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ is Euler's gamma function. If X_i 's are i.i.d. with a $\text{Gam}(\alpha, \beta)$ distribution, then the entropy of the standardized sum is

$$H \left(\frac{\beta}{\sqrt{n\alpha}} \sum_{i=1}^n \left(X_i - \frac{\alpha}{\beta} \right) \right) = H_g(n\alpha, \sqrt{n\alpha}).$$

With a slight abuse of notation let $H_g(\alpha) = H_g(\alpha, \sqrt{\alpha})$. It is easy to see that Theorem 2 below implies Theorem 1 if the common distribution of X_i belongs to the gamma family. \square

Theorem 2. The function $(1/2) \log(2\pi e) - H_g(\alpha)$ is completely monotonic on $\alpha \in (0, \infty)$.

Proof. Direct calculation gives

$$H_g(\alpha) = \log \Gamma(\alpha) + \alpha - \frac{1}{2} \log(\alpha) + (1 - \alpha) \psi(\alpha) \quad (2)$$

where $\psi(\alpha)$ is the digamma function defined by $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$. It is clear that $(1/2) \log(2\pi e) > H_g(\alpha)$, because the standard normal achieves maximum entropy among continuous distributions with unit variance. Moreover,

$$H'_g(\alpha) = (1 - \alpha) \psi'(\alpha) + 1 - \frac{1}{2\alpha}. \quad (3)$$

By Leibniz's rule, for $k \geq 2$,

$$(-1)^k H_g^{(k)}(\alpha) = (-1)^k \left[(1 - \alpha) \psi^{(k)}(\alpha) - (k - 1) \psi^{(k-1)}(\alpha) + \frac{(-1)^k (k - 1)!}{2\alpha^k} \right]. \quad (4)$$

The function $\psi^{(k)}(\alpha)$, $k \geq 1$, admits an integral representation (Abramowitz and Stegun, 1964, p. 260)

$$\psi^{(k)}(\alpha) = (-1)^{k+1} \int_0^\infty \frac{t^k e^{-\alpha t}}{1 - e^{-t}} dt. \quad (5)$$

Download English Version:

<https://daneshyari.com/en/article/1153468>

Download Persian Version:

<https://daneshyari.com/article/1153468>

[Daneshyari.com](https://daneshyari.com)