Original Research Paper

Whether Gaussian Nucleus Entropy Helps? Case in Point is Prediction of Number of Cesarean Births

Ramalingam Shanmugam

School of Health Administration, Texas State University-San Marcos, San Marcos, TX 78666, USA

Abstract: In this article, *entropy* in the collected data about the Gaussian population mean is traced from its embryonic stage as new data are periodically collected. The traditional *Shannon's entropy* has shortcomings from the data analytics point of view and it creates a necessity to refine the *Shannon's entropy*. Its refined version is named Gaussian *Nucleus Entropy* in this article. Advantages of the refined version are pointed out. The *Prior, likelihood, Posterior and predictive* nucleus entropies are derived, interconnected and interpreted. The results are illustrated using data on cesarean births in thirteen countries in the period [1987, 2007]. The medical communities and families are alarmed, as the cesarean births are increasing not due to emergency or necessity basis but rather for monetary or convenience basis. Nucleus entropy based data analysis answers whether their alarm is baseless.

Keywords: Prior, Posterior and Predictive Densities, Bayes Risk, Hypothesis Testing, Power

Introduction

What is entropy? In essence, entropy is all about information. Shannon (1948) introduced a seminal idea of capturing information in the collected data about an unknown parameter. In his example dealing with electronic communication, Shannon evaluated the amount of captured message in a destination location compared to the amount transmitted out of an originate location. Upon the advice of his friend John von Neumann, Shannon misnamed it as *entropy* for a reason. What was the reason? Claude Shannon himself revealed it: "My greatest concern was what to call it. I thought of calling it *information*, but the word was overly used, as I decided to call it *uncertainty*. When I discussed it with John von Neumann, he had a better idea. Von told me; you should call it *entropy* for two reasons. In the first place, your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, more important, *nobody knows what entropy really is*, so in a debate you will always have the advantage" (Tribus and McIrvine, 1971). However, the Shannon's entropy concept is popularly utilized in medicine, health, engineering, business, economics, marketing and statistics among other disciplines with a contextual interpretation opposite to what really, Shannon intended (Smelser and Baltes, 2001).

A litmus test for an entropy is that it should be quantifiable, partially orderable, additive, storable and transmittable. Shannon's entropy possesses several of those in the above list but does not the much-needed additive property, which is a requirement in data analysis. When an additional observation becomes available, the entropy ought to increase. The Shannon's entropy does not do so and hence, it needs a refinement. These and other controversies about the Shannon entropy convinced professionals to give up on entropy as documented in Ben-Naim (2011). Instead, why not we refine entropy and use it? Exactly, this question is answered in this article.

The Shannon's entropy is refined, in this article, by peeling away unnecessary entropy junkies. This refined new version is named *nucleus entropy*. Gaussian probability distribution is focused because it is more popularly employed in data analysis than any other distribution. Advantages in the *Gaussian nucleus entropy* over the traditional *Shannon's entropy* are pointed out.

Biostatistics community often selects Bayes approach over the frequentist's approach to perform data analysis and prediction as new data are periodically entering the databases. In the Bayes

approach, the *conjugate prior* jointly plays a cohesive crucial role along with that data dominant *likelihood function* to shape up the *posterior frequency density curve*. For sake of making a prediction of yet to be calculable, the *predictive frequency density curve* is often employed in the Bayesian analysis. In the course of time, a posterior frequency curve of the past becomes the prior for making the posterior frequency curve of the current time as new data become available. In this framework and practice, the Gaussian *conjugate prior, data dominant likelihood, posterior and predictive* nucleus entropies are derived, interconnected and interpreted.

In the illustration, the log-transformation of the number of cesarean births (in 1,000) over the years since 1987 in six groups of thirteen *countries: Belgium, Canada, Czech, Denmark, Finland, Ireland, Italy, Norway, Portugal, Slovak, Spain, Sweden and US* in Declercq *et al*. (2011) is considered, analyzed and interpreted. The future number of cesarean births in each of these thirteen countries is predicted and compared. Advantages of Gaussian nucleus entropy are articulated. A few conclusive comments are made in the end.

Gaussian Nucleus Entropies and Their Properties

To be specific, consider that a random sample $y_1, y_2, ..., y_n$ is drawn from a Gaussian population $f(y|\mu, \sigma_y^2)$ where Equation 1:

$$f\left(y|\mu, \sigma_y^2\right) = e^{-(y^2 - 2y\mu + \mu^2)/2\sigma_y^2} / \sqrt{2\pi\sigma_y^2};$$
$$-\infty < \mu < \infty; \sigma_y^2 > 0; -\infty < y < \infty, \quad (1)$$

μ is an unknown (natural) parameter and σ_y^2 is a known shape parameter. The Shannon's entropy $H_{Gauss}(.)$ is Equation 2a:

$$H_{Gauss}\left(y|\mu, \sigma_y^2\right)$$
$$= -\int f\left(y|\mu, \sigma_y^2\right) \ln f\left(y|\mu, \sigma_y^2\right) dy \quad (2a)$$
$$= -\frac{1}{2} \ln \sqrt{2\pi e \sigma_y^2}$$

Notice that its natural parameter μ is not even a part of the entropy (2). Furthermore, the Shannon's entropy exhibits a weakness from the practitioners' point of view. To see it, first note that it is known (Stuart and Ord, 1994) that a sum $s = y_1 + y_2 + ... + y_n$ of n independent and identically Gaussian distributed observations follows a *Gaussian probability structure*

$f(s|n\mu, n\sigma_y^2)$. In which case, the Shannon entropy of the sum $s = y_1 + y_2 + ... + y_n$ ought to be n times $H_{Gauss}(y|\mu, \sigma_y^2)$. However, it did not happen so. The Shannon's entropy of the sum is not additive of the individual entropies. That is:

$$H_{Gauss}\left(s|n\mu, n\sigma_y^2\right) \neq n H_{Gauss}\left(y|\mu, \sigma_y^2\right)$$

For entropy practitioners, this causes confusion, as the Shannon's entropy is not adding up as a new Gaussian observation becomes available. The existence of such a deficiency in Shannon's entropy is a sufficient reason to refine the entropy idea in an alternative way and it is what exactly done in this article. In other words, this article introduces a new and novel approach based on *Gaussian nucleus entropy*. That is:

Definition 1

A *nucleus entropy*, $N_{Gauss,Y}^{\mu, \sigma_y^2}$ resides in the Gaussian population frequency curve once it is written as Equation 2b:

$$f\left(y|\mu, \sigma_y^2\right) = A\left(\mu, \sigma_y^2\right) B\left(y, \sigma_y^2\right) N_{Gauss,Y}^{\mu, \sigma_y^2} \quad (2b)$$

With an *observation* $-\infty < y < \infty$, an unknown (*natural*) parameter $-\infty < \mu < \infty$ and an *entropy accumulator* parameter $\sigma_y^2 > 0$.

The function $A(\mu, \sigma_y^2)$ is insulated from observation y. The function $B(y, \sigma_y^2)$ does not connect to the unknown (natural) parameter μ. In a sense, both the functions (that is, $A(\mu, \sigma_y^2)$ and $B(y, \sigma_y^2)$) are *entropy junkies*. Unlike the Shannon's entropy, notice that the nucleus entropy involves both the unknown *natural* parameter μ and an *entropy accumulator* parameter σ_y^2. The Gaussian nucleus entropy is more appropriate, appealing and meaningful than the Shannon's entropy.

From now on, the variance σ_y^2 could perhaps be recognized and called as *nucleus entropy accumulator*. In other words, from (2b), the Gaussian nucleus entropy is

$N_{Gauss,Y}^{\mu, \sigma_y^2} = e^{\frac{y\mu}{\sigma_y^2}}$ with a single observation y and

$N_{Gauss,S}^{\mu, \sigma^2} = e^{\frac{s\mu}{\sigma_y^2}}$ for a sum $s = y_1 + y_2 + ... + y_n$ of n independent and identically distributed Gaussian observations.

Without losing any generality, the nucleus entropy could be expressed in a logarithmic scale just for the sake of a comparison with the Shannon's entropy, which is in a logarithmic scale. The expected *nucleus entropy* of a

single observation y is $E_{Gauss:Y}^{\mu,\sigma_y^2} = E_f\{\ln(N_{Gauss:Y}^{\mu,\sigma_y^2})\}$ which simplifies to Equation 3a:

$$E_{Gauss:Y}^{\mu,\sigma_y^2}$$
$$= \int\left[A(\mu,\sigma_y^2)B(y,\sigma_y^2)N_Y^{\mu,\sigma_y^2}\right]\ln N_Y^{\mu,\sigma_y^2} \qquad (3a)$$
$$= \frac{\mu^2}{\sigma_y^2}$$

The sample counterpart of (3a) is named *calculable Gaussian nucleus entropy* and it is Equation 3b:

$$O_{Gauss:Y}^{\mu,n} = \frac{\bar{y}^2}{\sigma_y^2} \qquad (3b)$$

The expected and observed *nucleus entropy* of a sum $s = y_1 + y_2 + \ldots + y_n$ of n random sample observations are $E_{Gauss:S}^{\mu,\sigma_y^2} = \frac{n\mu^2}{\sigma_y^2}$ and $O_{Gauss:S}^{\mu,\sigma_y^2} = \frac{n\bar{y}^2}{\sigma_y^2}$ respectively. In other words, the *expected and observed nucleus entropies of the likelihood function*:

$$f_{like}\left(y_1,y_2,\ldots y_n, \middle| \mu,\sigma_y^2\right) = e^{-\sum_{i=1}^n (y^2 - 2y\mu + \mu^2)/2\sigma_y^2} / \sqrt{2\pi\sigma_y^2}$$

Are respectively:

$$E_{Like,Gauss:S}^{\mu,\sigma_y^2} = \frac{n\mu^2}{\sigma_y^2} \qquad (4a)$$

And:

$$O_{Like,Gauss:S}^{\mu,\sigma_y^2} = \frac{n\bar{y}^2}{\sigma_y^2} \qquad (4b)$$

Notice that the *observed likelihood nucleus entropy (4b)* is calculable. The ratio of conditional expectation and variance:

$$Risk_{Like} = \frac{E\left(S\middle|\mu,\sigma_y^2\right)}{E\left(S\middle|\mu,\sigma_y^2\right) + var\left(S\middle|\mu,\sigma_y^2\right)} = \frac{\mu}{\left(\mu+\sigma_y^2\right)}$$

Is *explainable risk* in the likelihood to address the unknown Gaussian mean parameter μ (O'Hagan, 1994). Notice that the likelihood-based risk is more in the high neighborhood of the mean.

In the Bayesian analysis, the *prior distribution* for the unknown (natural) parameter of interest needs to be chosen in such a manner that the current data *likelihood function* is a dominant factor in the process of updating to obtain the *posterior distribution*. A *uniform* or a *conjugate prior* distribution is conventionally considered. The conjugate prior is more versatile than the uniform prior. In addition, in terms of the probability structure, the conjugate prior distribution is compatible with the likelihood function (Shanmugam, 1992). For the unknown (natural) Gaussian mean parameter, a Gaussian distribution Equation 5:

$$f_{prior}\left(\mu\middle|\bar{y}_0,\sigma_y^2/n_0\right) = \frac{\sqrt{n_0}}{\sqrt{2\pi\sigma^2}}e^{-n_0(\mu^2 - 2\bar{y}_0\mu + \bar{y}_0^2)/2\sigma_y^2},$$
$$-\infty < \mu < \infty; \sigma_y^2 > 0; -\infty < \bar{y}_0 < \infty, \qquad (5)$$

Is known (O'Hagan, 1994) to be conjugate prior distribution based on a prior sample of size $n_0 \geq 1$ and their mean $\bar{y}_0$. On its own merit, the conjugate prior distribution (5) is Gaussian and hence, the expected and observed prior Gaussian nucleus entropies are respectively:

$$E_{prior,Gauss:S}^{\mu,\sigma_y^2} = \frac{n_0\bar{y}_0^2}{\sigma_y^2} \qquad (6a)$$

And:

$$O_{prior,Gauss:S}^{\mu,\sigma_y^2} = \frac{n_0\mu^2}{\sigma_y^2} \qquad (6b)$$

A comment is necessary here and that the expected prior Gaussian nucleus entropy (6a) is calculable. Interestingly, there is a parallelism among expressions (4a), (4b), (6a) and (6b) and it is due to conjugation principle. The ratio of prior conditional expectation and variance:

$$Risk_{Prior} = \frac{E\left(\mu\middle|\bar{y}_0,\frac{\sigma_y^2}{n_0}\right)}{E\left(\mu\middle|\bar{y}_0,\frac{\sigma_y^2}{n_0}\right) + var\left(\mu\middle|\bar{y}_0,\frac{\sigma_y^2}{n_0}\right)} = \frac{n_0\bar{y}_0}{\left(n_0\bar{y}_0 + \sigma_y^2\right)}$$

Is *calculable prior risk* to address the unknown Gaussian mean parameter μ (O'Hagan, 1994). When the Gaussian nucleus entropy accumulation is zero in the beginning (that is, $\sigma_y^2 = 0$), the calculable *prior risk* is maximum possible one. Notice that the *calculable prior risk* reduces when the Gaussian nucleus entropy accumulation σ_y^2 increases. The importance of Gaussian nucleus entropy could not be clearer.

A beauty of the Bayesian analysis is that even if the prior distribution happens to be a bad selection, it is

eventually moderated in the posterior distribution. What is the posterior distribution? It is an update of the prior by its mixing with the data dominated likelihood function and integrating out their commonalities. Mathematically, it is done as follows:

$$
f_{posterior}\left(\mu \middle| \overline{y}_0, \overline{y}, \frac{\sigma_y^2}{n_0}, \frac{\sigma_y^2}{n}\right)
$$
$$
= \int_{-\infty}^{\infty} f_{prior}\left(\mu \middle| \overline{y}_0, \frac{\sigma_y^2}{n_0}\right) f_{like}\left(\overline{y} \middle| \mu, \frac{\sigma_y^2}{n}\right) d\overline{y} \qquad (7)
$$
$$
= \frac{\sqrt{(n+n_0)}}{\sqrt{2\pi\sigma_y^2}} e^{-(n+n_0)(\mu^2 - 2[n\overline{y}+n_0\overline{y}_0]\mu + [n\overline{y}+n_0\overline{y}_0]^2)/2\sigma_y^2},
$$
$$
-\infty < \mu < \infty; \sigma_y^2 > 0; -\infty < \overline{y}_0 < \infty, -\infty < \overline{y} < \infty.
$$

The posterior distribution (7) is also Gaussian and ence, the expected and observed posterior Gaussian nucleus entropies are respectively:

$$
E_{posterior, Gauss:S}^{\mu, \sigma_y^2} = \frac{(n\overline{y} + n_0\overline{y}_0)^2}{(n+n_0)\sigma_y^2} \qquad (8a)
$$

And:

$$
O_{prior, Gauss:S}^{\mu, \sigma_y^2} = \frac{(n+n_0)\mu^2}{\sigma_y^2} \qquad (8b)
$$

Again, the expected posterior Gaussian nucleus entropy (8a) is calculable. The ratio of posterior conditional expectation and variance:

$$
Risk_{Posterior}
$$
$$
= \frac{E\left(\mu \middle| \left(\frac{n\overline{y}+n_0\overline{y}_0}{n+n_0}\right), \frac{\sigma_y^2}{(n+n_0)}\right)}{E\left(\mu \middle| \left(\frac{n\overline{y}+n_0\overline{y}_0}{n+n_0}\right), \frac{\sigma_y^2}{(n+n_0)}\right) + var\left(\mu \middle| \left(\frac{n\overline{y}+n_0\overline{y}_0}{n+n_0}\right), \frac{\sigma_y^2}{(n+n_0)}\right)}
$$
$$
= \frac{(n\overline{y}+n_0\overline{y}_0)}{(n\overline{y}+n_0\overline{y}_0+\sigma_y^2)}
$$

Is *calculable posterior risk* to address the unknown Gaussian mean parameter μ (O'Hagan, 1994). When the Gaussian nucleus entropy accumulation is zero in the beginning (that is, $\sigma_y^2 = 0$), the calculable *posterior risk* is maximum possible one. Notice that the *calculable posterior risk* reduces when the Gaussian nucleus entropy accumulation σ_y^2 increases. The importance of Gaussian nucleus entropy cannot be over emphasized. The percent reduction in the *calculable risk* is:

$$
\Re_{calculable-risk}
$$
$$
= \left(Risk_{Posterior} - Risk_{Prior}\right) / Risk_{Prior} \qquad (8c)
$$
$$
= \left(\frac{n}{n_0}\right)\left(\frac{\overline{y}}{\overline{y}_0}\right)\left(\frac{\sigma_y^2}{\sigma_y^2 + n\overline{y} + n_0\overline{y}_0}\right)
$$

The purpose of knowledge discovery is only to reap its benefits currently or in the future. The Bayesian concepts meets this practical purpose. In other words, based on the prior or posterior distribution, the Bayesian approach leads into finding and using the so called *predictive distribution* to forecast what might be the future calculable mean, $\tilde{y}$ based on a yet to be collected random sample of size $m \geq 1$? The prior predictive density (O'Hagan, 1994) is Equation 9:

$$
\pi_{prior}(\tilde{y}) = \int_{-\infty}^{\infty} f_{prior}\left(\mu \middle| \overline{y}_0, \frac{\sigma_y^2}{n_0}\right) f_{like}\left(\tilde{y} \middle| \mu, \frac{\sigma_y^2}{m}\right) d\mu
$$
$$
= \frac{\sqrt{mn_0}}{\sqrt{2\pi[m+n_0]\sigma_y^2}} e^{-mn_0(\tilde{y}^2 - 2\overline{y}_0\tilde{y} + \overline{y}_0^2)/(2[m+n_0]\sigma_y^2)}, \qquad (9)
$$
$$
-\infty < \tilde{y} < \infty; \sigma_y^2 > 0; -\infty < \overline{y}_0 < \infty, m \geq 1
$$

Which is also Gaussian. Hence, the *expected and observed prior predictive Gaussian nucleus entropies* are respectively:

$$
E_{prior\,Predictive, Gauss:S}^{\mu, \sigma_y^2} = \frac{mn_0\overline{y}_0^2}{(m+n_0)\sigma_y^2} \qquad (10a)
$$

And:

$$
O_{prior\,Predictive, Gauss:S}^{\mu, \sigma_y^2} = \frac{mn_0\tilde{y}\overline{y}_0}{(m+n_0)\sigma_y^2} \qquad (10b)
$$

Again, the expected posterior Gaussian nucleus entropy (10a) is calculable.

Due to the posterior distribution (7), the posterior predictive density is Equation 11:

$$
\pi_{posterior}(\tilde{y})
$$
$$
= \int_{-\infty}^{\infty} f_{posterior}\left(\mu \middle| \left[\frac{n\overline{y}+n_0\overline{y}_0}{n+n_0}\right], \frac{\sigma_y^2}{(n+n_0)}\right) f_{like}\left(\tilde{y} \middle| \mu, \frac{\sigma_y^2}{m}\right) d\mu
$$
$$
= \frac{\sqrt{m(n+n_0)}}{\sqrt{2\pi[m+n+n_0]\sigma_y^2}} \qquad (11)
$$
$$
e^{-\frac{m(n+n_0)}{(2[m+n+n_0]\sigma_y^2)}(\tilde{y}^2 - 2\tilde{y}[\frac{n\overline{y}+n_0\overline{y}_0}{n+n_0}] + [\frac{n\overline{y}+n_0\overline{y}_0}{n+n_0}]^2)},
$$
$$
-\infty < \tilde{y} < \infty; \sigma_y^2 > 0; -\infty < \overline{y}_0, \overline{y} < \infty, m, n, n_0 \geq 1
$$

Which is also Gaussian. Hence, the *expected and observed posterior predictive Gaussian nucleus entropies* are respectively Equation 12a and b:

$$E_{posterior \, \text{Pr} edictive, Gauss:S}^{\mu, \sigma_y^2}$$
$$= \frac{m(n+n_0)\left[\dfrac{n\bar{y}+n_0\bar{y}_0}{n+n_0}\right]^2}{(m+n+n_0)\sigma_y^2} \tag{12a}$$

And:

$$O_{posterior \, \text{Pr} edictive, Gauss:S}^{\mu, \sigma_y^2}$$
$$= \frac{m(n+n_0)\tilde{y}\left[\dfrac{n\bar{y}+n_0\bar{y}_0}{n+n_0}\right]}{(m+n+n_0)\sigma_y^2} \tag{12b}$$

Again, the expected posterior Gaussian nucleus entropy (10a) is calculable. The extra *quadratic information*: $\mathbb{Q}_{\inf ormation}$ in the current sample compared to what it was in the prior sample is Equation 13:

$$\mathbb{Q}_{\inf ormation}$$
$$= var\left(\mu \middle| \bar{y}_0, \frac{\sigma_y^2}{n_0}\right) - var\left(\mu \middle| \left(\frac{n\bar{y}+n_0\bar{y}_0}{n+n_0}\right), \frac{\sigma_y^2}{(n+n_0)}\right) \tag{13}$$
$$= \left(\frac{n}{n+n_0}\right)\sigma_y^2$$

Which suggests that the extra quadratic information increases when the Gaussian nucleus entropy accumulation increases and/or the current sample size far exceeds the prior sample size.

The *calculable jump* in the predictive Gaussian nucleus entropy from prior to posterior time is Equation 14:

$$\mathfrak{I}_{predictiveGauss:S}^{\sigma_y^2} = \frac{E_{posterior \, \text{Pr} edictive, Gauss:S}^{\mu, \sigma_y^2}}{E_{prior \, \text{Pr} edictive, Gauss:S}^{\mu, \sigma_y^2}}$$
$$= \left(\frac{n_0}{n+n_0}\right)\left(\frac{m+n_0}{m+n+n_0}\right)\left(\frac{n_0\bar{y}_0+n\bar{y}}{n_0\bar{y}_0}\right)^2. \tag{14}$$

The *calculable jump* from the prior Gaussian nucleus entropy to the posterior nucleus entropy is Equation 15:

$$\wp_{Gauss:S}^{\sigma_y^2} = \frac{E_{posterior, Gauss:S}^{\mu, \sigma_y^2}}{E_{prior, Gauss:S}^{\mu, \sigma_y^2}}$$
$$= \left(\frac{n_0}{n+n_0}\right)\left(\frac{n_0\bar{y}_0+n\bar{y}}{n_0\bar{y}_0}\right)^2. \tag{15}$$

The proportionality of predictive entropy jump $\mathfrak{I}_{predictiveGauss:S}^{\sigma_y^2}$ per unit of jump of $\wp_{Gauss:S}^{\sigma_y^2}$ is Equation 16:

$$\mathfrak{R}_{predictiveGauss:S}^{\sigma_y^2}$$
$$= \frac{\mathfrak{I}_{predictiveGauss:S}^{\sigma_y^2}}{\wp_{predictiveGauss:S}^{\sigma_y^2}} \tag{16}$$
$$= \left(\frac{m+n_0}{m+n+n_0}\right)\left(\frac{n}{n+n_0}\right)\left(\frac{n_0}{n+n_0}\right)\left(\frac{O_{Like, Gauss:S}^{\mu, \sigma_y^2}}{O_{Like, Gauss:S_0}^{\mu, \sigma_y^2}}\right)$$

where, $O_{Like, Gauss:S}^{\mu, \sigma_y^2}$ and $O_{Like, Gauss:S}^{\mu, \sigma_y^2}$ denote the Gaussian likelihood nucleus entropy (4b) of the current sample sum $s = y_1 + y_2 + ... + y_n$ and the prior sample sum $s_0 = y_1 + y_2 + ... + y_{n_0}$. Hence, we obtain the Theorem 1 below.

Theorem 1

Because the first three factors in (16) are less than one, when:

$$\left(\frac{O_{Like, Gauss:S}^{\mu, \sigma_y^2}}{O_{Like, Gauss:S_0}^{\mu, \sigma_y^2}}\right) = \left(\frac{n}{n_0}\right)\left(\frac{\bar{y}}{\bar{y}_0}\right)^2 < 1$$

The Gaussian nucleus entropy in the current sample is lesser than its counterpart in the prior sample and hence, the current sample is *lesser informative* than the prior sample.

When the Theorem 1 prevails, further sampling might not be worthwhile.

Case in Point is Gaussian Nucleus Entropy to Predict Number of Cesarean Births

In this section, the results of the previous section are illustrated using the data on the number Y of cesarean births (in 1,000 s) incountries: *Belgium, Canada, Czech, Denmark, Finland, Ireland, Italy, Norway, Portugal, Slovak, Spain, Sweden and US* during 1987 through 2002 in Declercq *et al.* (2011). The first and foremost data analysis checks whether the $\ln Y$ (because the numbers are big in size) follows a Gaussian distribution. Figure 1 to 5 and they confirm that the $\ln Y$ (because the dots are closer to the upward diagonal line) indeed follows a Gaussian distribution. Hence, the data are suitable for the illustration of the Gaussian nucleus entropy results of the previous section. Figure 6 to realize the proximity and trend among the years of the incidences.

In our analysis, we evaluate the data before the year 2,000 as prior sample (with $n_0 = 3$) and the data after the year 2,000 as current sample (with $n = 2$). Table 1 for the values of $\ln Y, \bar{y}_0, \bar{y}, \sigma_y^2$ and the observed likelihood nucleus

entropy $O_{Like,Gauss:S}^{\mu,\sigma_y^2}$ in (4b) for the entire sample. The expected prior nucleus entropy $E_{prior,Gauss:S}^{\mu,\sigma_y^2}$ in (6a), the expected posterior nucleus entropy $E_{posterior,Gauss:S}^{\mu,\sigma_y^2}$ in (8a), the calculable risk $\Re_{calculable-risk}$ in (8c), the quadratic information $\mathbb{Q}_{\inf ormation}$ in (13), The *calculable jump* in the predictive Gaussian nucleus entropy from prior to posterior time $\mathfrak{I}_{predictiveGauss:S}^{\sigma_y^2}$ in (14), The *calculable jump* from the prior Gaussian nucleus entropy to the posterior nucleus entropy is $\wp_{Gauss:S}^{\sigma_y^2}$ in (15) and the proportionality $\Re_{predictiveGauss:S}^{\sigma_y^2}$ of predictive entropy jump $\mathfrak{I}_{predictiveGauss:S}^{\sigma_y^2}$ per unit of jump of $\wp_{Gauss:S}^{\sigma_y^2}$ in (16) of the current over prior sample likelihood entropies are calculated and displayed in the Table 2.
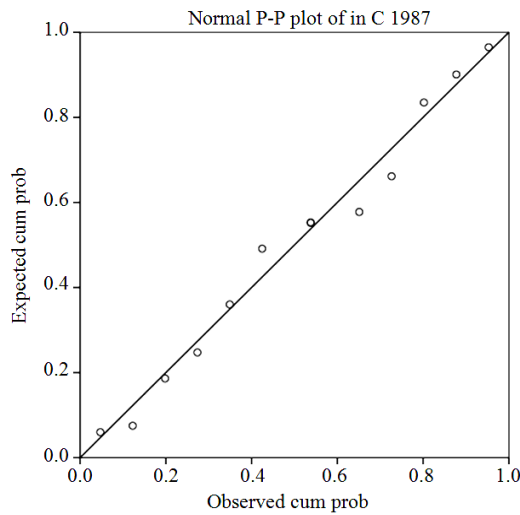

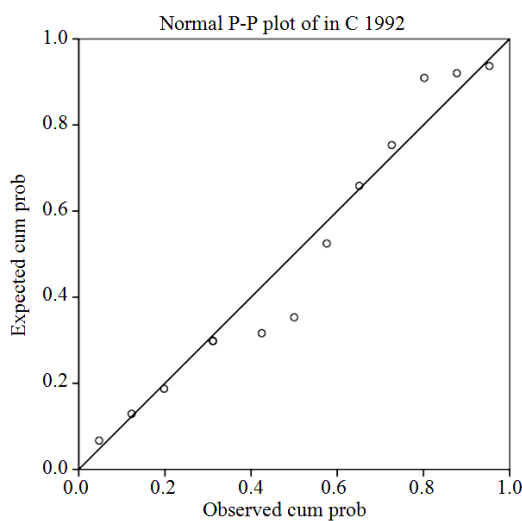Fig. 3. Gaussian of Cesarean births in 1997
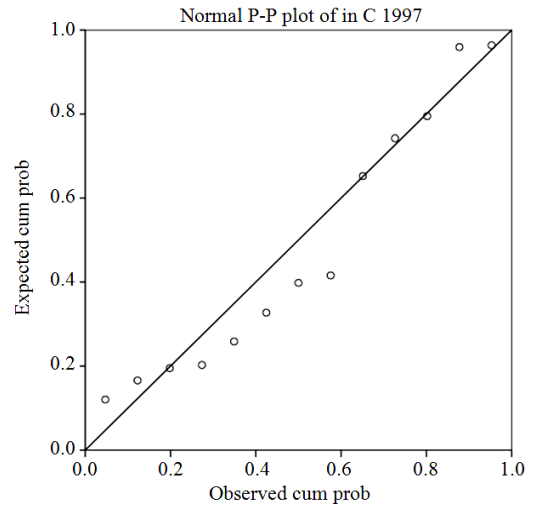

Fig. 1. Gaussian of cesarean births in 1987

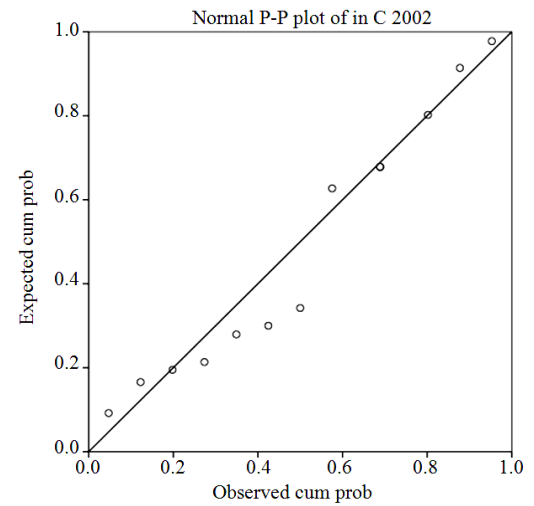Fig. 4. Gaussian of Cesarean births in 2002

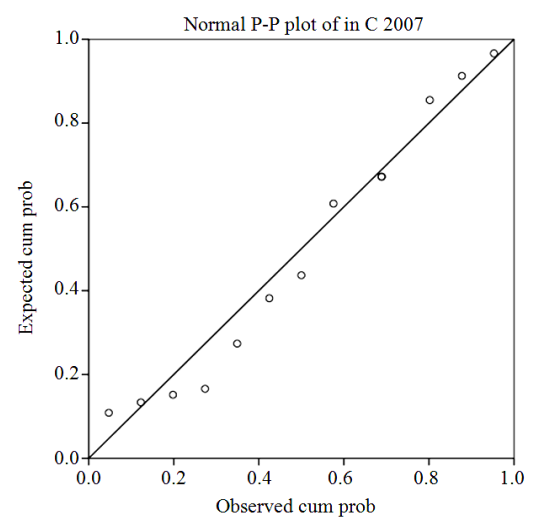Fig. 2. Gaussian of cesarean births in 1992

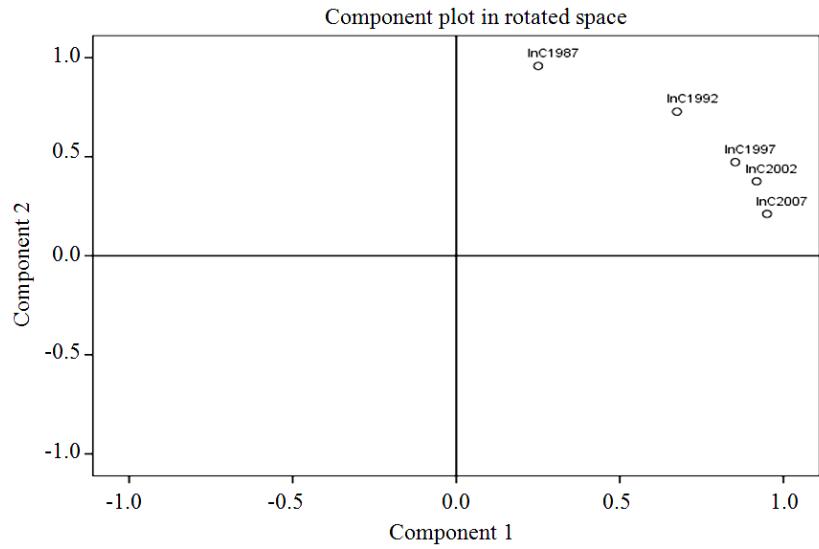

Fig. 5. Gaussian of cesarean births in 2007

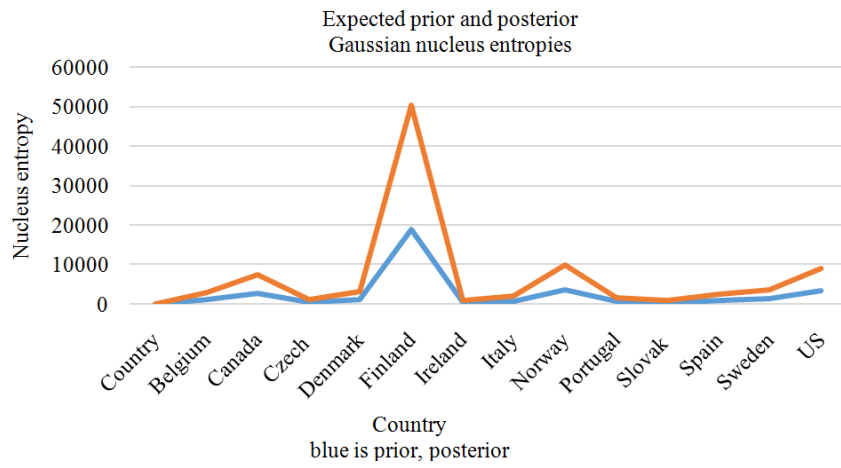Fig. 6. Proximity among the years as far as the incidence of cesarean births

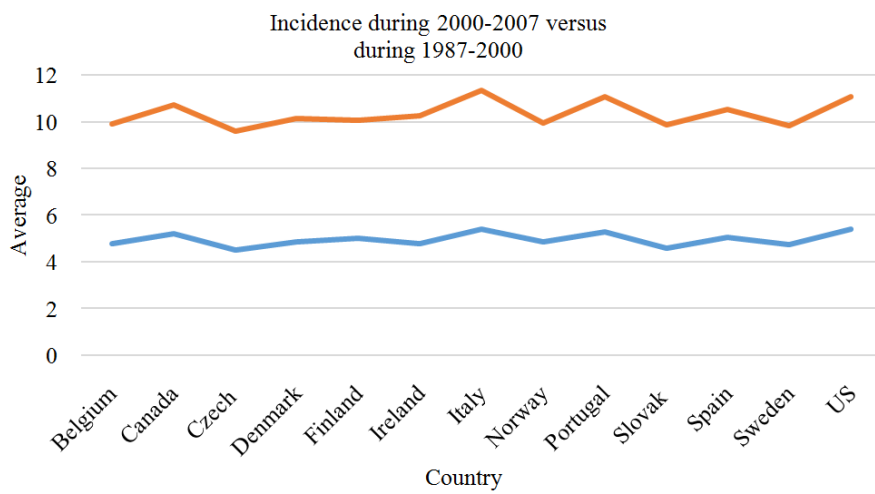Fig. 7. Expected prior and posterior Gaussian nucleus entropy for the countries

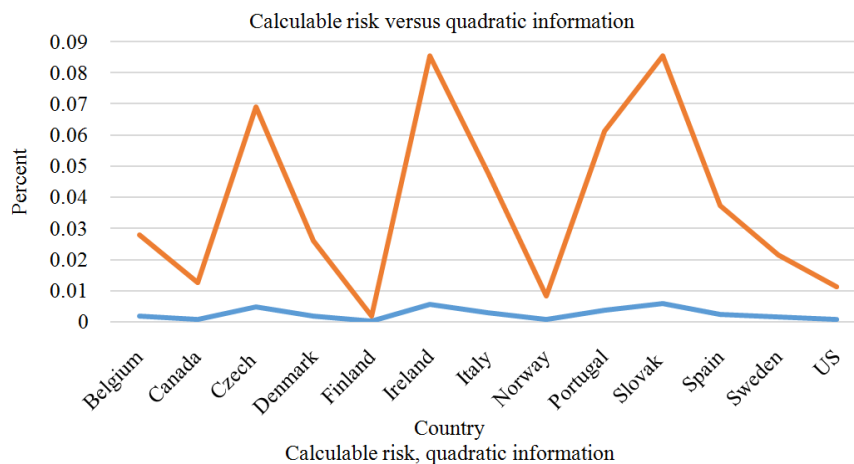Fig. 8. Cesarean births have increased after 2,000

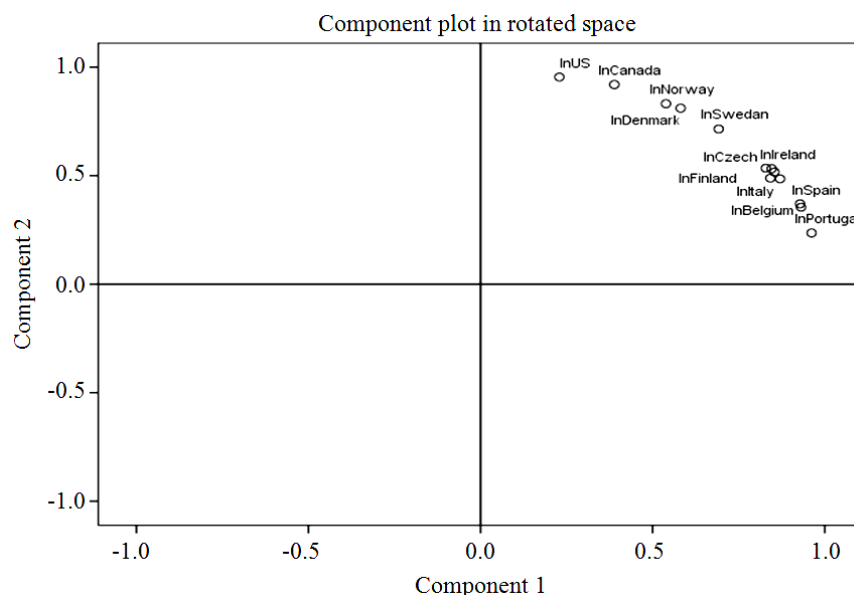Fig. 9. Calculable risk versus Quadratic information for the countries

Fig. 10. Proximity among the countries

Table 1. The average and variance of ln (# Cesarean Births) in 1,000 since 1987

	Before 2,000			After 2,000		Prior mean 4.78	Sample mean 5.16	variance 0.07	Posterior mean 4.93	Observed nucleus 1866.1
Country	1987	1992	1997	2002	2007					
Belgium	4.56	4.79	4.98	5.16	5.15	5.22	5.52	0.03	5.34	4833.50
Canada	5.28	5.18	5.21	5.46	5.58	4.5	5.12	0.16	4.75	699.39
Czech	4.25	4.48	4.77	4.95	5.28	4.86	5.29	0.06	5.03	2093.70
Denmark	4.86	4.80	4.92	5.21	5.37	4.99	5.10	0.00	5.03	31836.00
Finland	4.96	4.98	5.04	5.10	5.09	4.77	5.50	0.20	5.06	641.65
Ireland	4.49	4.79	5.03	5.42	5.58	5.40	5.94	0.11	5.62	1400.60
Italy	5.16	5.45	5.60	5.89	5.99	4.85	5.11	0.02	4.95	6331.80
Norway	4.86	4.84	4.86	5.08	5.13	5.29	5.78	0.14	5.49	1044.10
Portugal	4.88	5.38	5.61	5.71	5.85	4.59	5.30	0.20	4.87	595.15
Slovak	4.29	4.60	4.87	5.18	5.41	5.06	5.50	0.09	5.23	1561.30
Spain	4.80	5.09	5.29	5.46	5.53	4.73	5.11	0.05	4.88	2386.20
Sweden	4.68	4.67	4.83	5.05	5.17	5.41	5.66	0.03	5.51	5752.40
US	5.47	5.41	5.34	5.56	5.76	4.78	5.16	0.07	4.93	1866.10

Table 2. Summary of expected prior posterior, predictive, likelihood entropies, risk and information

Country	$E^{\mu,\sigma_y^2}_{prior,Gauss:S}$	$E^{\mu,\sigma_y^2}_{posterior,Gauss:S}$	$\Re_{calculable-risk}$	$\mathbb{Q}_{inf\,ormation}$	$\Im^{\sigma_y^2}_{predictiveGauss:S}$	$\wp^{\sigma_y^2}_{Gauss:S}$	$\dfrac{O^{\mu,\sigma_y^2}_{Like,Gauss:S}}{O^{\mu,\sigma_y^2}_{Like,Gauss:S_0}}$
Belgium	1051.9380	1797.9490	0.00185	0.02603	1.367	1.709	0.710
Canada	2772.6840	4705.5100	0.00076	0.01181	1.358	1.697	0.697
Czech	377.2589	656.5635	0.00489	0.06441	1.392	1.740	0.742
Denmark	1171.8010	2008.6880	0.00169	0.02419	1.371	1.714	0.715
Finland	18794.0000	31527.7500	0.00011	0.00159	1.342	1.678	0.678
Ireland	341.8576	598.0071	0.00573	0.07987	1.399	1.749	0.751
Italy	777.3856	1337.1490	0.00283	0.04507	1.376	1.720	0.721
Norway	3646.2610	6178.3590	0.00054	0.00775	1.356	1.694	0.695
Portugal	582.4762	999.7821	0.00369	0.05765	1.373	1.716	0.717
Slovak	316.7505	554.3287	0.00595	0.07970	1.400	1.750	0.752
Spain	875.5363	1499.6460	0.00235	0.03509	1.370	1.713	0.713
Sweden	1343.1690	2297.0890	0.00143	0.01996	1.368	1.710	0.711
US	3325.6020	5626.0880	0.00066	0.01055	1.353	1.692	0.692

The Fig. 7 confirms that the expected posterior Gaussian nucleus entropy dominates the corresponding expected prior Gaussian nucleus entropy in every country. This finding would have been missed if the Shannon's entropy is used and it emphasizes the importance of the nucleus entropy of this article. First, the Fig. 8 informs that the average incidence of cesarean births during the years 2,000-2,007 has been consistently more than the average incidence of cesarean births during the years 1,987-2,000 in every country. It, therefore, answers that the concern of the medical professionals and the communities of increasing cesarean births is legitimate. The extra *quadratic information*: $\mathbb{Q}_{inf\,ormation}$ in the current sample compared to what it was in the prior sample is consistently more than the percent reduction in the *calculable risk* $\Re_{calculable-risk}$ in each country (Fig. 9) and it confirms that the data information is more than the risk of addressing the Gaussian population mean, according to the nucleus entropy. Such confirmation is not possible when the Shannon's entropy is involved and it demonstrates the importance of nucleus entropy as a refinement of the Shannon's entropy. Finally, we notice (Fig. 10) that the countries cluster together with respect to all of the above mentioned nucleus entropy related results. In specific, the countries: *US and Canada* form the first cluster, the countries: *Norway, Sweden and Denmark* form the second cluster and the countries: *Czech, Finland, Ireland, Italy, Spain, Belgium and Portugal* form the third cluster. The visualization of such clusters of countries is feasible due to the concept of nucleus entropy and it is not possible in using the complicated Shannon's entropy.

Conclusion

In conclusion, the central theme, nucleus entropy of this particle refines the seminal and wonderful Shannon's concept to capture data information. Though this refined nucleus entropy has been illustrated using the cesarean data, it is versatile enough to be useful in finance, economics, marketing, engineering, medical and health studies also. Further research on building regression models using nucleus entropy concept is underway and it will be communicated later in a journal article.

Author's Contributions

There is no coauthor in this work and hence, all contributions were from the author Dr. Ram Shanmugam.

Ethics

The author, Dr. Ram Shanmugam takes responsibility for any error or opinion in the article.

References

Ben-Naim, A., 2011. A Farewell to Entropy: Statistical Thermodynamics Based on Information. 1st Edn., World Scientific Press, Singapore.

Declercq, E.L., R. Young, H. Cabral and J. Ecker, 2011. Is a rising cesarean delivery rate inevitable? Trends in industrialized countries, 1987 to 2007. Birth, 38: 99-104. PMID: 21599731

O'Hagan, A., 1994. Kendall's Advanced Theory of Statistics, Bayesian Inference. 1st Edn., Wiley, New York, ISBN-210: 0470233818, pp: 300.

Smelser, N.J. and P.B. Baltes, 2001. International Encyclopedia of the Social and Behavioral Sciences. 1st Edn., Elsevier, ISBN-10: 0080430767, pp: 733.

Shanmugam, R., 1992. Quantifying Prior Opinion in Length Biased Mean Natural Exponential Family. In: Bayesian Analysis in Statistics and Econometrics, Goel, P.K. and N.S. Iyengar (Eds.), Springer, New York, ISBN-13: 978-1-4612-2944-5, pp: 389-395.

Shannon, C.E.A., 1948. A mathematical theory of communication. Bell Sys. Tech. J., 27: 379-423. DOI: 10.1002/j.1538-7305.1948.tb01338.x

Stuart, A. and J. Ord, 1994. Kendall's Advanced Theory of Statistics. 1st Edn., Griffin Publications, London.

Tribus, M. and E.C. McIrvine, 1971. Energy and information. Scientific Am., 224: 179-186.