World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:3, No:9, 2009

The Possibility-Probability Relationship for Bloodstream Concentrations of Physiologically Active Substances

Arkady Bolotin

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

***Abstract*—**If a possibility distribution and a probability distribution are describing values x of one and the same system or process $x(t)$, can they relate to each other? Though in general the possibility and probability distributions might be not connected at all, we can assume that in some particular cases there is an association linked them.

In the presented paper, we consider distributions of bloodstream concentrations of physiologically active substances and propose that the probability to observe a concentration x of a substance *X* can be produced from the possibility of the event $X = x$.

The proposed assumptions and resulted theoretical distributions are tested against the data obtained from various panel studies of the bloodstream concentrations of the different physiologically active substances in patients and healthy adults as well.

***Keywords*—**Possibility distributions, possibility-probability relationship.

I. Introduction

FOR many years, the modeling of uncertainty in life sciences has been totally dominated by probabilistic methods. Now, following major criticism of the principal limitation of probability theory, other approaches to dealing with uncertainty are being explored [1, 10].

The principal limitation we are talking about consists in the probability theory's inability to address the reality of partiality – partial truth, partial precision, and partial possibility. Therefore, standard probabilistic methods cannot reflect one of the main features of the life sciences systems [12].

Let us take a few simple examples showing off this fact.

1. Imagine we are asked the question: What is the probability that a given newborn will have anemia? The difficulty of using probability theory here is rooted in the basic property of conditional probabilities: given $P(X)$ all that can be said about $P(X|Y)$ is that its value is between zero and one. Accordingly, if we were told that 28% of Jewish babies born to families settled in the Negev had anemia, it would tell nothing about the probability that the given baby would have anemia. The same holds true even if we got more information about this particular newborn. Say this baby is Jewish born in the Negev, then all that we would be able to know is the fraction (i.e. 0.28) of babies having anemia in the particular category, but all that can be said about the probability that this baby will have anemia is that it is between 0 and 1.

2. Suppose we have a group of *N* adults and we are asked: What is the fraction K/N of them that have the systolic blood pressure lying in the interval 115±2 (mmHg)? Obviously, all we can say is that this fraction K/N is somewhere between zero and one. Next, assume we are told that *M* individuals of those *N* are young men, and we are asked again: What is the fraction L/M of these men that have the systolic blood pressure lying in the interval 115±2? Would be this fraction L/M higher than the fraction K/N ? Again, we cannot answer these questions. All we can say is that the value L/M is between 0 and 1. As in the previous example, additional information does not improve our ability to estimate the probability of the systolic blood pressure equal to 115±2.

It is worth to mention here, that unlike probability theory, in *possibility theory*, learning more about the process $x(t)$ means restricting the range of possible values for x. (In fact, possibility distributions hold *negative* information: they do not support but exclude facts [5, 11].) This means in particular, that additional information does improve our ability to estimate the possibility distribution: for example, the possibility of the sys-

Author is with Epidemiology Department, Ben-Gurion University of the Negev, Beersheba 84105, Israel (e-mail: arkadyv@bgu.ac.il).

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

tolic blood pressure equal to 115±2 should be higher in the group of *M* young men than in the general group of *N* individuals.

Consequently, possibility theory seems to be one of the most promising mathematical concepts in life sciences, especially for modeling complex biological or medical processes.

On the other hand, in modeling uncertainty with possibility theory the foundational question is how to obtain a possibility distribution for a given system/process. The obvious way – to extract a possibility distribution from the experiment by measurements – needs for the analysis of the relation between possibility and probability distributions. Such analysis is essential because unlike probability distributions *possibility distributions cannot be directly measured* [9].

Though in general the possibility and probability distributions describing the same system/process might be not connected at all, we can assume that in some particular cases there is an association linked them.

In our previous paper [4], we build the possibility distributions for the controlled bloodstream concentrations x of some physiologically active substances through few approximate considerations. These possibility distributions were then tested against the empirical histograms obtained from the panel study of the eight different physiologically active substances in 417 individuals.

In that paper, we did not analyze the exact relation between the resulted possibility distributions $\mu(x)$ and probability distributions $p(x)$ but merely put forward the simple hypothesis that the possibility distribution $\mu(x)$ can predict with fair certainty *the trend* for the probability distribution: $p(x) \square \mu(x)$.

Now, in the present paper, we intend to analyze this relation in detail.

II. Building the Possibility Distribution through Approximate Considerations

It is known that the bloodstream concentration of every physiologically active substance (such as a hormone, a protein, a steroid, a triglyceride, a mineral or a trace mineral) is controlled by at least two processes: *secretion* and *utilization*. The process of secretion is responsible for production of the substance and its release into the blood stream, while the utilization process removes the substance from the blood through consumption or degradation [7, 8].

Let the *continuous* and *finite* function $\mu(x)$ (together with its derivatives of at least the first and second orders) be the possibility distribution of a given substance X taking concentrations x in bloodstream.

Considering the assumption that the interval $[0;\infty)$ contains all possible concentrations x, there must be at least one *impossible* concentration $x' \in [0;\infty)$ such that $\mu(x') = 0$ and one *possible without any restriction* concentration $x'' \in [0;\infty)$ such that $\mu(x'') = 1$; thus, the following expression must be true: $\mu(x)\big|_{0 \le x < \infty} \in [0,1]$.

Though for different substances the processes of secretion and utilization may vary in their particular realization, we believe that the following three general assumptions hold for all of them:

Assumption 1. If the substance X is controlled by the secretion-utilization processes, its bloodstream level x can never drop to zero. In other words, zero concentration $x = 0$ of the controlled physiologically active substance X is *impossible*:

$$\mu_X(0) = 0 \quad . \tag{1}$$

Assumption 2. A very high level of the substance X controlled by the secretion-utilization processes is impossible too, but not quite much as zero level. It is so because the precise highest limit $x_{\max}$ for bloodstream concentrations x of a given substance does not exist. Therefore, the possibility of high concentrations x must vanish only *asymptotically*:

$$\mu_X(x)\big|_{x \to \infty} = 0 \quad . \tag{2}$$

Assumption 3. For each substance X controlled by the secretion-utilization processes, it must be the level of equilibrium x_0 between these processes, that the concentration x_0 of the substance X is *possible without any restriction*. The further from the equilibrium x_0 in either side, the less the concentration x is possible:

$$\mu_X(x_0) = \max\left\{\mu_X(x)\big|_{0 \le x < \infty}\right\} = 1 \quad . \tag{3}$$

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

Since the function $\mu_X(x)$ must be continuous, from the *Assumption 1* it follows that $\mu_X(x)$ must be defined at every point near zero:

$$0 = \mu_X(0) = \mu_X(+0) \quad . \tag{4}$$

Hence, near zero the function $\mu_X(x)$ should take the form:

$$\mu_X(0)\big|_{x\to 0} \sim x^m, \quad m > 0 \quad . \tag{5}$$

The only function that can vanish asymptotically while be finite at zero is an exponential function of a negative argument; so the *Assumption 2* can be written down as follows:

$$\mu_X(x)\big|_{x\to\infty} \sim \exp\left(-\alpha x^s\right), \quad \alpha > 0, \quad s > 0 \quad . \tag{6}$$

Thus, the function $\mu_X(x)$ describing the possibility distribution of the concentrations x must take the form

$$\mu_X(x) = Cx^m \exp\left(-\alpha x^s\right) \quad . \tag{7}$$

We have postulated that the function $\mu_X(x)$ is finite together with its derivatives of the first and the second orders; this postulate can fit the conjecture (7) if only we have

$$\begin{cases} m = 1+n & , \quad n = 0 \text{ or } n \geq 1 \\ s = (+0)^n + k & , \quad k \geq 0 \end{cases} \quad . \tag{8}$$

According to the *Assumption 3*, the function $\mu_X(x)$ has the maximum at x_0, this gives us the value of the parameter α

$$\alpha = ms^{-1}x_0^{-1} \tag{9}$$

and the constant C

$$C = x_0^{-m} \exp\left(\frac{m}{s}\right) \quad . \tag{10}$$

Consequently, the function $\mu_X(x)$ takes the form

$$\mu_X(z;n;k) = z^{1+n} \cdot \exp\left\{\frac{1+n}{(+0)^n + k} \cdot \left[1 - z^{(+0)^n + k}\right]\right\} \quad , \tag{11}$$

where z denotes *dimensionless concentrations* of the substance X

$$z = \frac{x}{x_0} \quad . \tag{12}$$

As it can be readily seen from the Eq. (11), the only parameter of the function $\mu_X(z;n;k)$ that cannot be equal to zero is the equilibrium concentration x_0. Therefore, the "simplest" form of the Eq. (11) is the following one:

$$\mu_X(z;0;0) = z \cdot \exp(1-z). \tag{13}$$

We propose that the function (13) represents the possibility distribution of the bloodstream concentrations for each physiologically active substance controlled by the secretion-utilization processes.

III. Relation between the Possibility and Probability Distributions

Let the continuous and finite (together with its derivatives of at least the first and second orders) function $p_X(x)$ be the probability distribution of the bloodstream concentrations x of a given substance X controlled by the secretion-utilization processes.

Regarding the relationship between the distributions $\mu_X(x)$ and $p_X(x)$ describing the same substance X, we put forward the following assumptions:

Assumption 4. The probability distribution $p_X(x)$ can be produced from the possibility distribution $\mu_X(x)$.

Assumption 5. If the concentration x is impossible, it cannot be probable either. The opposite is not true: the concentration x may not be probable but still possible.

Assumption 6. If the concentration x is possible without any restriction, it is also the most probable one.

Mathematically, the *Assumption 4* signifies this:

$$p_X(z) = F\left[\mu_X(z)\right] \quad . \tag{14}$$

From the *Assumption 5* it follows

$$0 \leq \left.\frac{F\left[\mu_X(z)\right]}{\mu_X(z)}\right|_{0\leq z<\infty} < \infty \quad , \tag{15}$$

and the *Assumption 6* means that

$$\left.\frac{F\left[\mu_X(z)\right]}{\mu_X(z)}\right|_{z=1} = \text{const} > 0 \quad . \tag{16}$$

The Eq. (15) and (16) do take place if $F[\mu_X(z)]$ is the exponential function of the $\mu_X(z)$:

$$p_X(z) = Cz^m \cdot \exp(m - mz) \quad , \qquad (17)$$

where the exponent *m* is equal to or greater than 1, and *C* is a constant.

Let us assume that this is really so and find out the *m*.

IV. Statistical Estimation of the Exponent *m*

To do this, we will try to estimate the best polynomial model of the kernel density $k_X(z)$ – used as an approximation of the probability distribution $p_X(z)$ – among the following 13 models of a given bloodstream substance *X*:

$$k_X(z) = \left[\mu_X(z)\right]^n + \text{const}, \quad n = -2, -1\frac{1}{2}, -1, \ldots, 4 \quad . \qquad (18)$$

We will use the data obtained from various panel studies of the bloodstream concentrations of the different physiologically active substances in patients and healthy adults as well done by the scientists and health professionals of the Ben-Gurion University of the Negev in the period from 1997 to 2006 years [2, 3, 6].

The Table I shows the best power *n* among the 13 models defining by the Eq. (18) for each studied substance.

TABLE I
The Best Polynomial Model of the Kernel Density $k_x(z)$ on the Possibility Distribution $\mu_x(z)$

Substance	*N* of observations	Mode concentration x_0		Best power *n*
Serum ferritin	744	26.5±8.8	ng/mL	3.5
Serum iron	1001	37.0±9.2	μg/dL	2
Zinc	441	102.0±28.0	μmol/L	4
Albumin	75	5.0±4.4	U/L	3
Glucose	633	126.0±19.2	mg/dL	4
Triglyceride	629	163.0±52.2	mg/dL	1.5

The standard deviation of each mode concentration x_0 was calculated by using the 50 bootstrap replications of the given x_0.

As it follows from this table, the average best power *n* across the analyzed substances is three. This gives us the reason to believe that the exponent *m* in the Eq. (17) is equal to three too.

Consequentially, we get the formula for the probability distribution $p_X(z)$.

$$p_X(z) = \frac{27}{2} \cdot z^3 e^{-3z} \quad . \qquad (19)$$

In the following figures, we present the theoretical distribution (19) together with the observed histograms of the studied substances.

V. Tests of the Equality of the Empirical and Theoretical Distributions

According to the Eq. (19), the theoretical mean concentration $\overline{z}_{\text{theor.}}$ is equal to:

$$\overline{z}_{\text{theor.}} = \frac{27}{2}\int_0^\infty z^4 e^{-3z} dz = \frac{4}{3} = 1.3(3) \qquad (20)$$

and the theoretical standard deviation $\sigma_{\text{theor.}}$ of the concentrations *z* is:

$$\sigma_{\text{theor.}} = \sqrt[2]{\frac{27}{2}\int_0^\infty (z - \overline{z})^2 \cdot z^3 e^{-3z} dz} = \frac{2}{3} = 0.6(6) \qquad (21)$$

As it follows from the data presented in the Tables II and III, the observed mean concentrations $\overline{z}$ and their standard deviations σ do not statistically differ from the corresponding theoretical values (20) and (21).

TABLE II
The Empirical and Theoretical Means

Substance	Observed means $\overline{z}$	95% CI for $\overline{z}$	$\overline{z}_{\text{theor.}}$	Probability that $\overline{z} = \overline{z}_{\text{theor.}}$
Serum ferritin	1.35	0.89÷1.78	1.3(3)	> 0.05
Serum iron	1.35	1.00÷1.66		> 0.05
Zinc	1.33	0.97÷1.70		> 0.05
Albumin	1.54	0.15÷2.51		> 0.05
Glucose	1.28	1.13÷1.54		> 0.05
Triglyceride	1.08	0.91÷1.76		> 0.05

TABLE III
The Empirical and Theoretical Standard Deviations

Substance	Observed σ	95% CI for σ	$\sigma_{\text{theor.}}$	Probability that $\sigma = \sigma_{\text{theor.}}$
Serum ferritin	0.86	0.45÷0.89	0.6(6)	> 0.05
Serum iron	0.69	0.50÷0.83		> 0.05
Zinc	0.65	0.48÷0.85		> 0.05
Albumin	1.20	0.08÷1.26		> 0.05
Glucose	0.48	0.57÷0.77		> 0.05
Triglyceride	0.61	0.45÷0.88		> 0.05

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

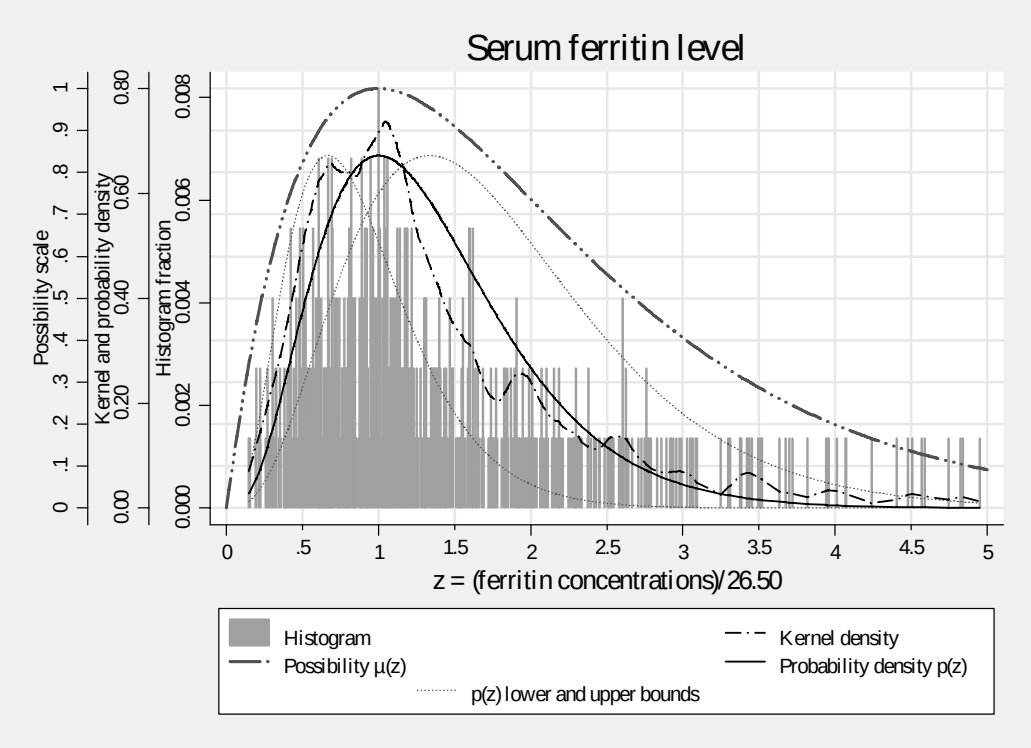

Fig. 1 The histogram and the theoretical distributions for serum ferritin level

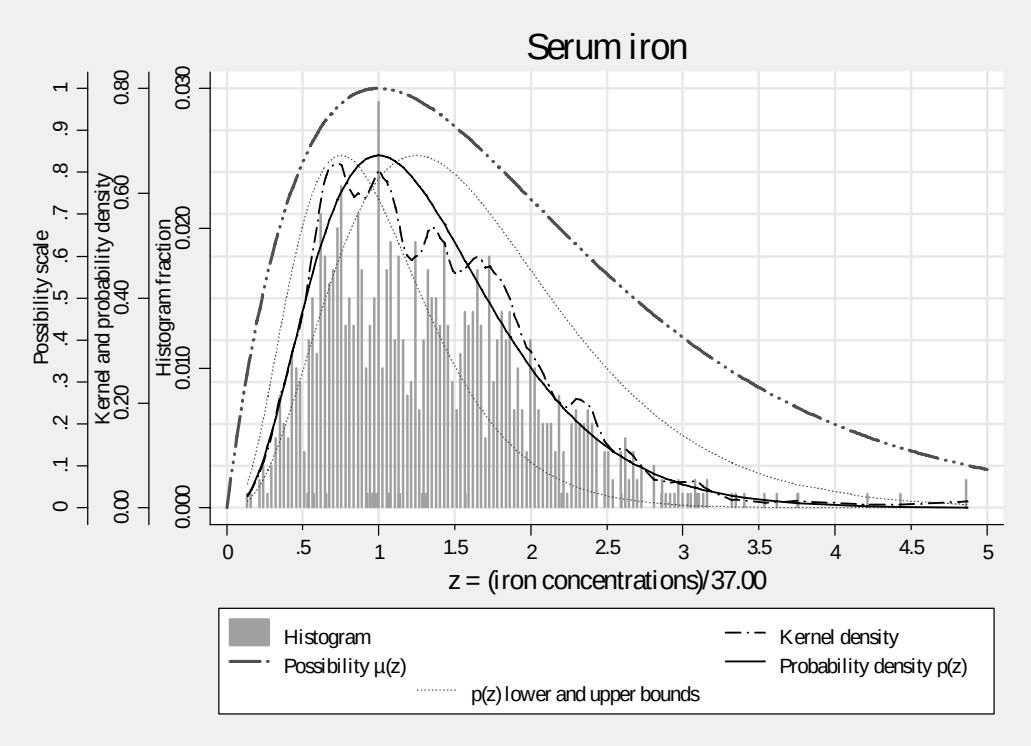

Fig. 2 The histogram and the theoretical distributions for serum iron

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

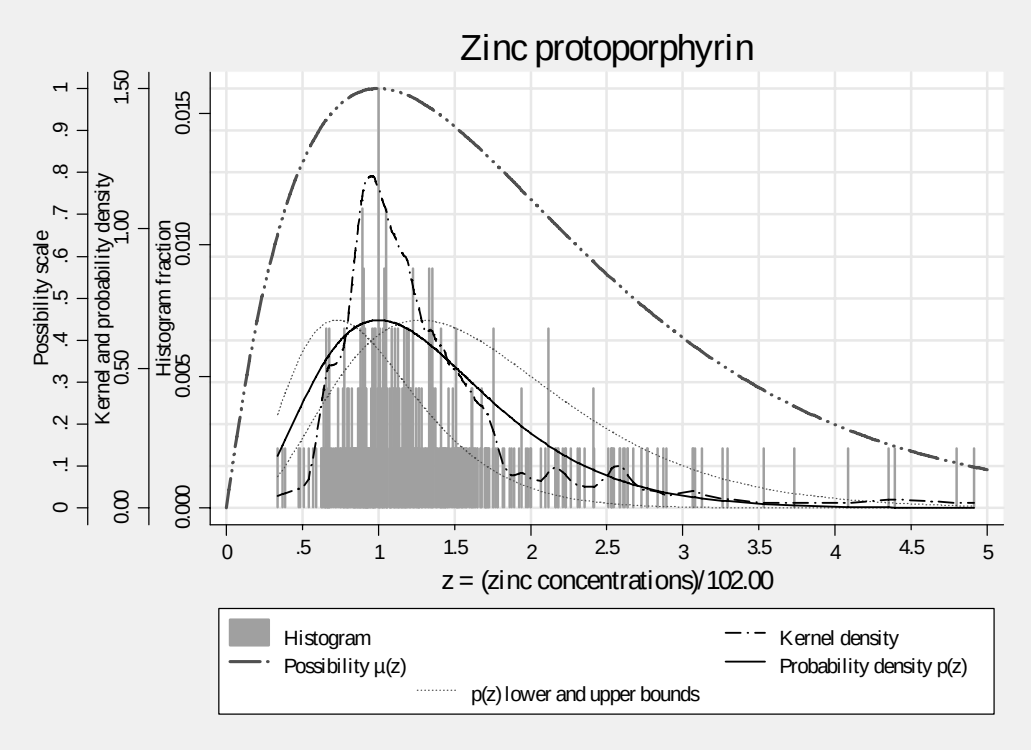

Fig. 3 The histogram and the theoretical distributions for zinc protoporphyrin

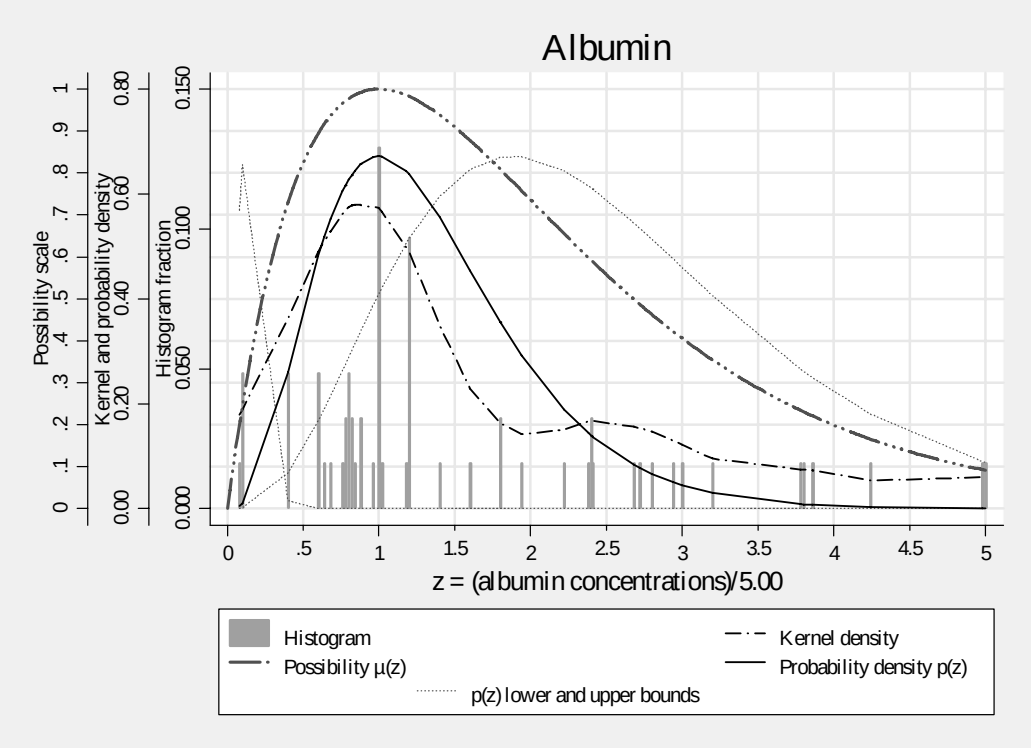

Fig. 4 The histogram and the theoretical distributions for albumin

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

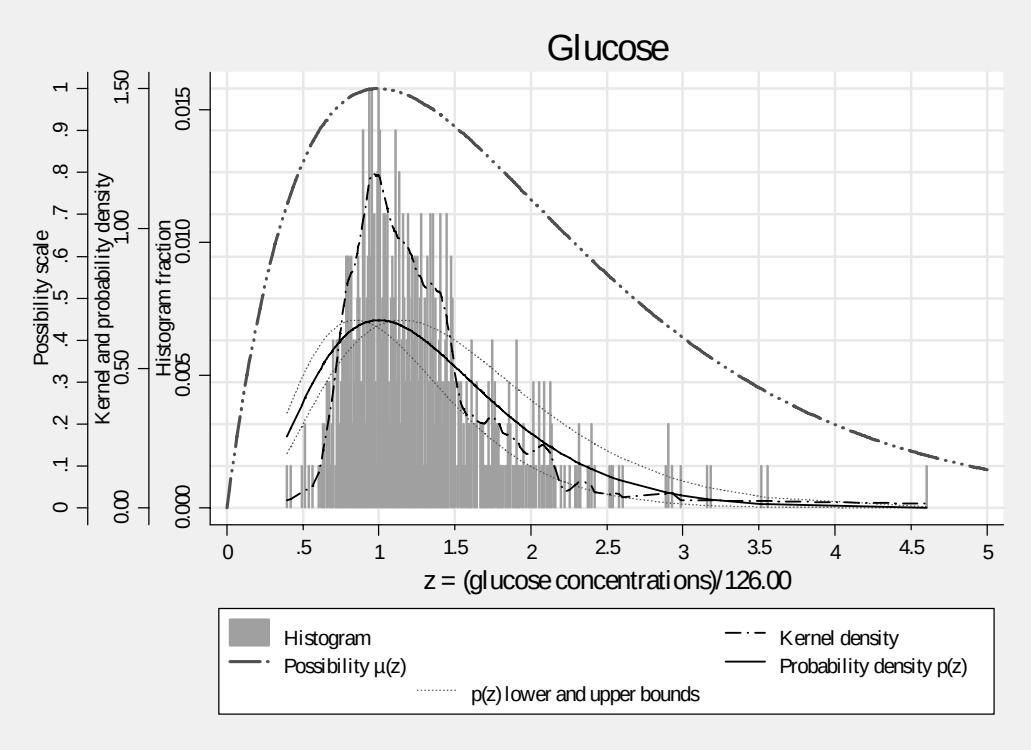

Fig. 5 The histogram and the theoretical distributions for glucose

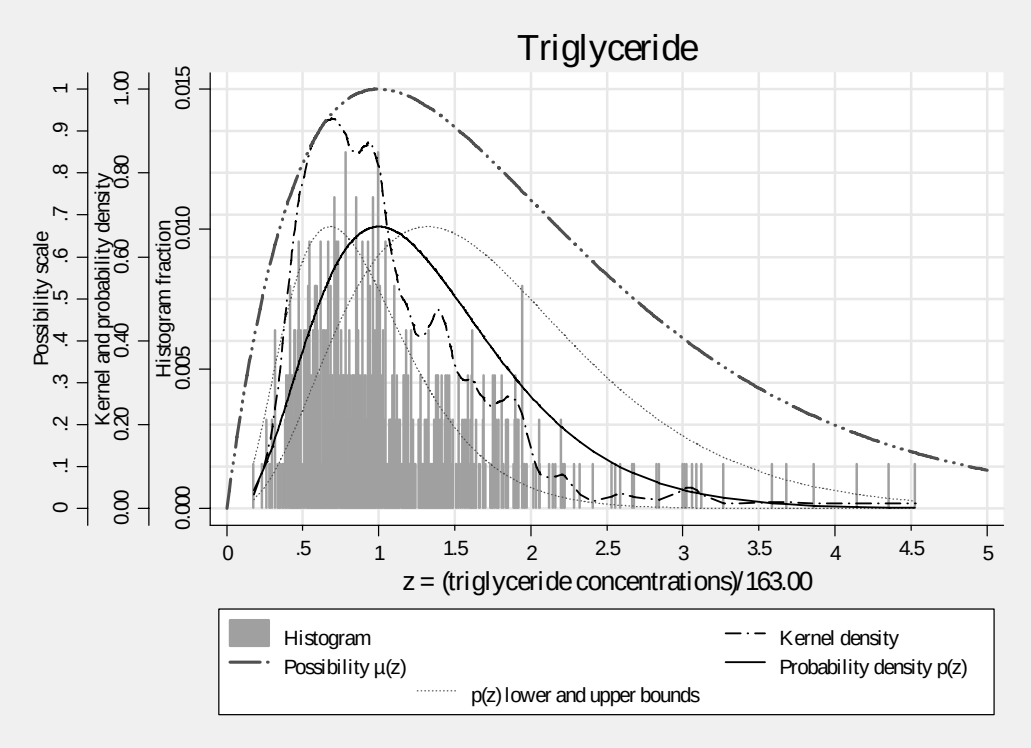

Fig. 6 The histogram and the theoretical distributions for triglyceride

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804

These results unequivocally support the assumptions we made about the functional relationship between the possibility distribution $\mu_X(z)$ and the probability distribution $p_X(z)$ for the bloodstream concentrations z of physiologically active substances.

REFERENCES

[1] Akay, M., Cohen, M., Hudson, D. Fuzzy sets in life sciences. Fuzzy Sets and Systems, 90, 1997, pp. 219-224.
[2] Bilenko N, Yehiel M, Inbar Y, Gazala E. The association between anemia in infants, and maternal knowledge and adherence to iron supplementation in southern Israel. The Israel Medical Association journal : IMAJ, 2007, 9(7), pp. 521-524.
[3] Bilenko, N., Shahar, D.; Shai, I.; Weitzman, S.; Fraser, D. Prevalence and characteristics of myocardial infarction, dia-betes and hypertension in the adult Jewish population: re-sults from the Negev Nutritional Study. Harefuah, 142(1), 2003, pp. 17-21.
[4] Bolotin, A. The Possibility Distribution for the Controlled Bloodstream Concentrations of Any Physiologically Active Substance. PWASET Vol. 23, 2007, ISSN 1307-6884, pp. 61-66.
[5] Dubois, D. and Prade, H. Fuzzy sets in approximate reason-ing, Part 1: Inference with possibility distributions. Fuzzy Sets and Systems, Vol. 40, 1991, pp. 143-202.
[6] Fraser D, Bilenko N, Vardy H, Abu-Saad K, Shai I, Abu-Shareb H, Shahar DR. Differences in food intake and dispar-ity in obesity rates between adult Jews and Bedouins in southern Israel. Ethnicity & disease, 2008, 18(1), pp. 13-18.
[7] Mathews, C. K. and Holde, K. E. Integration and control of metabolic processes. In: D. Bowen. Biochemistry. s.l. : Ben-jamin/Cummings Publishing Group, 1990, pp. 790-792.
[8] Medical Encyclopedia. Medline Plus. [Online] A service of the U.S. National Labrary of Medicine and the National In-stitutes of Health, Date last updated: 26 February 2009. [Cited: December 11, 2009.] http://www.nlm.nih.gov/medlineplus/encyclopedia.html.
[9] Nguyen, H.T. Fuzzy sets and probability. Fuzzy Sets and Systems, 90, 1997, pp. 129-132.
[10] Rouvray, D. H. The treatment of uncertainty in the sciences. Endeavour, Vol. 21 (4), 1997, pp.154-158.
[11] Spott, M. A theory of possibility distributions. Fuzzy Sets and Systems, Vol. 102, 1999, pp. 135-155.
[12] Zadeh, L. A. Toward a perception-based theory of probabil-istic reasoning. Journal of Statistical Planning and Inference, 105, 2002, pp. 233–264.

Open Science Index, Medical and Health Sciences Vol:3, No:9, 2009 waset.org/Publication/2804