Possibility distribution: a unified representation for parameter estimation

Gilles P. Mauris

LISTIC, Polytech'Savoie, Université de Savoie
Annecy, France
Email: mauris@univ-savoie.fr

Abstract— *the paper presents a possibility formulation of one-parameter estimation that unifies some usual probability formulations. Point and confidence interval estimation are unified in a single theoretical formulation and incorporated into estimators of an omnibus form: a possibility distribution. New relationships between continuous possibility distribution and probability concepts are established. The concept of specificity ordering of a possibility distribution is then used for comparing the efficiency of different estimators. The usefulness of the approach is illustrated on mean and median estimators from data sample of different size and of different probability distribution.*

Keywords— efficiency, parameter estimation, possibility, probability, stochastic ordering, uncertainty.

1 Introduction

Systems designed by engineers are often meant to influence their environment: to manipulate it, to move it, to control it and so on. To enable such actions, these systems need information, e.g. values of physical quantities describing their environment. Generally, two types of information sources are available: prior knowledge and empirical knowledge. The latter is obtained by sensor observations. Prior knowledge is the knowledge that was already there before a given observation became available. The combination of prior knowledge and empirical knowledge leads to posterior knowledge useful for the task at hand. Parameter Estimation deals with this problem to infer a parametric description for an object, a physical process or an event, given measurements that always come with uncertainties due to variability of influence quantities or phenomena. Therefore an assessment of the parameter estimate uncertainty is also required by the application, and prior knowledge is often used to reduce the uncertainty of the final estimate. To tackle such problems a lot of developments have been done in the statistical science [1].

In statistics, a parameter estimator $T_\theta(X)$ is defined as a function of the observable sample data $X_1, X_2, ..., X_n$ that is used to estimate an unknown parameter θ; an estimate $\hat{\theta}$ is the result from the actual application of the function to a particular set of data. Prior knowledge can be that $X_1, X_2, ..., X_n$ is an independent identical distributed (*iid*) sample from a continuous random variable X having a distribution function F. Many different estimators are possible for any given parameter. Two main types of estimators are worldwide used: point estimators and interval estimators. The selection of an estimator of one of the above kinds for purposes of informative inference, including typical applications in scientific research is generally admitted to involve elements of choice which are in some degree arbitrary. Such elements include the choice of a particular confidence level for an interval estimator, and the choice of a location function for a point estimator for a given confidence coefficient. In addition, a point estimate is often desired along with an interval. Such considerations and related ones have led to proposals for simultaneously use of a point confidence estimator and a set of confidence limit or interval estimation having various confidence coefficients [2]. Such estimators may be regarded as a modern formulation of a long standing practice of reporting estimates in the form $\hat{\theta} \pm k\sigma$, where k is some constant and σ the standard deviation of the observations. Note that this expression is recommended by the Guide of Uncertainty in Measurement (GUM [3]) edited by international legal metrology institutions. This form may be interpreted as an ordered set of three point estimators. Tukey proposed in [4] that for general purposes, it would be advantageous to use a set of five points estimators at standard confidence level *2,5%,16,66%, 50%, 83,33% and 97,5%*. Cox proposed to use the full continuous family of confidence levels varying between *0* and *1* by introducing a so called confidence curve estimator [5].

The problem of choosing a good estimator, i.e. an estimator which tends to take values close to the true unknown value, is formulated mathematical by introducing criteria of closeness. The latter has been specified by the introduction of specific loss function (called sometimes risk function). For example, the absolute error criterion was introduced by Laplace, Gauss replaced this by the squared error criterion which proved mathematically much more tractable [2]. Each such definite specification of closeness can be criticized as somewhat arbitrary, except in a context when one postulates the reality of the indicated cost of errors of each possible kind. This point put forth by Galileo leads to comparisons between estimators on the basis of all their probabilities rather than on the basis of certain summaries such as the variance [2]. Birnbaum extended this concept to continuous random variables by introducing the notion of peakedness of a probability distribution, which characterizes the concentration of values around the estimate [6].

ISBN: 978-989-95079-6-8

In this paper, we propose a possibility formulation of one parameter estimation, which unifies the usual previous probability formulation. Our proposition is based on a possibility/probability transformation we previously proposed [7][8], and on the notion of possibility distribution specificity ordering [9]. In section 2, this transformation is recalled and new relationships with the above mentioned probability estimation concepts are exhibited. In particular, the efficiency of estimators is characterized by its stochastic concentration defined by the possibility specificity. The approach is illustrated in section 3 on the problems of mean and median estimation. The advantages of the possibility formulation, which leads to an intuitive graphic representation, are highlighted, particularly when the sample size varies and the probability distribution as well.

2 Possibility versus probability estimation

2.1 Basics of possibility theory

Possibility theory is one of the modern theories available to represent uncertainty when information is scarce and/or imprecise [10]. The basic notion is the possibility distribution [11], denoted π. Here, we consider possibility distributions defined on the real line, i.e. π is an upper semi-continuous mapping from the real line to the unit interval. Thus π is a fuzzy subset but with specific semantics for the membership function. Indeed, a possibility distribution describes the more or less plausible values of some uncertain variable *X*. The possibility theory provides two evaluations of the likelihood of an event, for instance whether the value of a real variable *X* does lie within a certain interval: the possibility Π and the necessity N. The normalized measures of possibility Π and necessity N are defined from the possibility distribution $\pi : R \to [0,1]$ such that $\sup_{x\in R} \pi(x) = 1$ as follows:

$\forall A \subset R, \Pi(A) = \sup_{x\in A} \pi(x)$ and

$\forall A \subset R, N(A) = 1 - \Pi(\overline{A}) = \inf_{x\notin A}(1-\pi(x))$

The possibility measure Π satisfies:

$\forall A, B \subset R, \Pi(A \cup B) = \max(\Pi(A), \Pi(B))$

The necessity measure N satisfies:

$\forall A, B \subset R, N(A \cap B) = \min(N(A), N(B))$

In fact, possibility measures are set functions similar to probability measures, but they rely on axioms which involve the operations “ maximum “ and “ minimum “ instead of the operations “addition” and “product” (if the measures are decomposable).

Moreover we can interpret any pair of dual functions necessity/possibility $[N, \Pi]$ as upper and lower probabilities induced from specific probability families. Let π be a possibility distribution inducing a pair of functions $[N, \Pi]$. We define the probability family $\mathcal{P}(\pi) = \{P, \forall A, N(A) \le P(A)\} = \{P, \forall A, P(A) \le \Pi(A)\}$. In this case, $\sup_{P\in\mathcal{P}(\pi)} P(A) = \Pi(A)$ and $\inf_{P\in\mathcal{P}(\pi)} P(A) = N(A)$ [12]. In other words, the family $\mathcal{P}(\pi)$ is entirely determined by the probability intervals it generates. In the estimation context, this fact can be useful when it is not possible to identify one specific probability distribution for the observations. The concept of specificity can be used to qualify the informativeness of a possibility distribution. Indeed, a possibility distribution π_1 is said more specific than π_2 as soon as $\forall x, \pi_1(x) \le \pi_2(x)$ [9](it is the usual definition of inclusion of fuzzy sets), i.e. π_1 is more informative than π_2.

2.2 Background on probability/possibility transformation

Let us assume that the sample data are issued from a continuous random variable X defined on the set of reals and described by a probability density function f, F being its corresponding cumulative distribution function, with F^{-1} its inverse function if it exists (otherwise the pseudo-inverse function can be considered [13]). For every possible confidence level $\beta \in [0,1]$, the corresponding confidence interval is defined as an interval that contains the parameter to be estimated, i.e. θ, with probability $\ge \beta$. In other words, a confidence interval of confidence level β (denoted I_β) is defined as an interval for which the probability P_{out} to be outside this interval I_β does not exceed $\alpha \overset{def}{=} 1-\beta$, i.e. $P(\theta \notin I_\beta) = \alpha$.

It is possible to link confidence intervals and possibility distribution in the following way. A unimodal numerical possibility distribution may be viewed as a nested set of confidence intervals, which are the α cuts $[\underline{x}_\alpha, \overline{x}_\alpha] = \{x, \pi(x) \ge \alpha\}$ of π. The degree of certainty that $[\underline{x}_\alpha, \overline{x}_\alpha]$ contains μ is $N([\underline{x}_\alpha, \overline{x}_\alpha])$ (if π continuous). Obviously, the confidence intervals built around the same point x^* are nested. It has been proven in [7] that stacking confidence intervals of a probability distribution on top of one another leads to a possibility distribution (denoted π^* having x^* as modal value). In fact, in this way, the α-cuts of π^*, i.e. $A_\alpha = \{x \mid \pi^*(x) \ge \alpha\}$ are identified with the confidence interval I_β^* of confidence level $\beta = 1-\alpha$ around the nominal value x^*. Thus, the possibility distribution π^* encodes the whole set of confidence intervals in its membership function. Moreover, this possibility distribution satisfies:

$\forall A \subset R, \Pi^*(A) \ge P(A)$, with Π^* and P the possibility and probability measures associated respectively to π^* and f.

A closed form expression of the possibility distribution $\pi^M(x)$ induced by confidence intervals around the median $x^* = M$ is obtained for symmetric unimodal continuous probability densities $f(x)$ strictly increasing on the left and decreasing on the right of M [7]:

$$\begin{aligned} &\forall x \in [-\infty, M], \pi^M(x) = 2F(x) \\ &\forall x \in [M, +\infty], \pi^M(x) = 2(1-F(x)) \end{aligned} \qquad (1)$$

Thus, the possibility distribution is closely related to the cumulative distribution function.

ISBN: 978-989-95079-6-8

2.3 Relationships with probability estimation notions

According to the way a possibility distribution is built from the probability distribution F from which the sample data come from, the link with confidence intervals is immediate. Indeed, generally the parameter to be estimated is a location parameter of a distribution function, e.g. the mean. The link with Cox confidence curve $c_\theta(x)$ is also obvious. Indeed the latter is defined by [5]:

$\forall\theta\in[-\infty,x], c_\theta(x)=F(\theta-x)$

$\forall\theta\in[x,+\infty], c_\theta(x)=1-F(\theta-x)$

Thus, $c_\theta(x)=\pi_F(x-\theta)/2$.

Other new links concern the quality of estimator that is related to the concentration of values around the estimate, quantified for example by the variance or by the absolute mean deviation. Hereafter a few proposition formalizing for continuous symmetric random variables the links with the absolute mean deviation, the Levy concentration function, and other stochastic orderings.

Proposition 1:

The specificity index $sp(\pi)=\int_{-\infty}^{+\infty}\pi^M(x)dx$ of the possibility $\pi^M(x)$ distribution equivalent to a continuous symmetric random variable X is equal to twice the absolute mean deviation:

$$sp(\pi_X^M)=2.E\left|X-M\right|.$$

Proof:

Assume without loss of generality that *M=0*. Due to symmetry:

$$\int_{-\infty}^{+\infty}\pi_X^M(x)dx=\int_{-\infty}^{0}\pi_X^M(x)dx+\int_{0}^{+\infty}\pi_X^M(x)dx=$$

$$2\int_{-\infty}^{0}\pi_X^M(x)dx=4\int_{-\infty}^{O}F(x)dx$$

In other respect, as $F(-\infty)=0$:

$$2\int_{-\infty}^{0}F(x)dx=2\left[xF(x\right]_{-\infty}^{0}-2\int_{-\infty}^{0}xf(x)dx$$

$$=0-\int_{-\infty}^{\infty}xf(x)dx=E(\left|X\right|)$$

Therefore: $sp(\pi_X^M)=\int_{-\infty}^{+\infty}\pi_X^M(x)dx=2.E\left|X-M\right|$.

Thus the possibility distribution is closely related to the absolute mean deviation, (and not to the conventional standard deviation), of course if the latter is defined. Indeed, when the term $\int_{-\infty}^{0}xf(x)dx$ is infinite, the possibility distribution can be computed but the associated specificity is infinite.

Proposition 2:

The possibility distribution $\pi_X^M(x)$ equivalent to a continuous unimodal symmetric random variable X is related to the Lévy concentration function $Q_X(x)$ by the relation:

$$\forall x\geq 0, Q_X(x)=1-\pi_X^M(x+M)$$

Proof:

The concentration function introduced by P. Lévy in 1935 has been first defined by the expression [14, page 44]:

$\forall x\geq 0, Q_X(x)=\sup_{x_0}\left[F(x_0+x)-F(x_0-x)\right]$

Note that, in a latter definition [14, page 90], which is the one considered usually, the expression has became: $Q_X(x)=\sup_{x_0}\left[F(x_0+x)-F(x_0)\right]$. Here we will consider the first definition, more close to our approach.

If F is a symmetric and unimodal distribution about M then, the supremum in equation is attained for $x_0=M$ (see [8] for a demonstration). Thus we have:

$\sup_{x_0}\left[F(x_0+x)-F(x_0-x)\right]=F(M+x)-F(M-x)$.

The symmetry of F gives for $x\geq 0$:

$$\begin{aligned}&F(M+x)-F(M-x)=F(M+x)-(1-F(M+x))\\&=2F(M+x)-1=1-2(1-F(M+x))=1-\pi(x)\end{aligned}.$$

Therefore, $Q_X(x)=1-\pi_X^M(x+M)$.

In fact, the possibility distribution gathers probability dispersion intervals and it is complementary to the Lévy concentration function for continuous unimodal random variables.

Proposition 3:

The specificity order (defined by the fuzzy subset inclusion) defines a stochastic concentration ordering equivalent for continuous unimodal symmetric random variables both to the peakedness ordering of Birnbaum [6] and to the majorization ordering of Hickey [15]:

$$\pi_X^M(x)\leq\pi_Y^M(x)\Leftrightarrow X\geq^{peaked}Y\Leftrightarrow X\leq^{maj}Y$$

Proof:

According to Birnbaum [6], X is said to be more peaked about θ, than Y about θ if and only if:

$\Pr(\left|X-\theta\right|\geq t)\leq\Pr(\left|Y-\theta\right|\geq t)$ holds for all $t\geq 0$.

It is clear that we have:

$\pi_Y(\theta-t)=\pi_Y(\theta+t)=1-\Pr(\theta-t,\theta+t)=\Pr(\left|X-\theta\right|\geq t$

Thus: $\pi_X^M(x)\leq\pi_Y^M(x)\Leftrightarrow X\geq^{peaked}Y$

Note also that peakenedness ordering is equivalent to conventional stochastic ordering of absolute variables:

$$X\geq^{peaked}Y\Leftrightarrow\left|X-\theta\right|\leq^{sto}\left|Y-\theta\right|\overset{def}{\Leftrightarrow}F_{\left|X-\theta\right|}(x)\leq F_{\left|Y-\theta\right|}(x)$$

The majorization order introduced by Hickey is used to compare continuous distribution in terms of randomness. Y is said at least as random as X in the majorization sense if [15]:

$$\int_o^t g^*(y)dy\leq\int_o^t f^*(y)dy,\forall t,$$

where f^* and g^* are the decreasing rearrangement of f and g respectively; that is:

$f^*(x)=\sup\{c:m(c)>x\}, x>0$ with,

$m(c)=\mu\{x:f(x)>c\}$, μ denoting Lebesgue measure and f the corresponding density function of F.

ISBN: 978-989-95079-6-8

It is proved in [16] that for unimodal continuous random variables:

$$\int_o^t f^*(y)dy = Q_X(t) .$$

Thus: $\pi_X^M(x) \le \pi_Y^M(x) \Leftrightarrow Q_X(x) \ge Q_Y(x) \Leftrightarrow X \le^{maj} Y$.

Note also that a similar proposition as the proposition 3, but for discrete probability and possibility distributions has been stated by Dubois and Hüllermeier [17].

Thus given two estimators $T_\theta(X)$ and $U_\theta(X)$ of a parameter θ, T is said to have greater stochastic concentration about than U does if T is more peaked than U. An interesting consequence of greater peakedness proved by Hwang [18] is that if L is any loss function L satisfying: $L(x,\theta) = h(|x-\theta|)$,

where for each θ, $h:[0,\infty) \to [0,\infty)$ is a non decreasing function , we have:

$$\pi_{T_\theta(X)}(x) \le \pi_{U_\theta(X)}(x) \Leftrightarrow E_\vartheta(L(T_\theta(X)) \le E_\vartheta(L(U_\theta(X))$$

In other words not only is T at least as likely as U to be within a neighborhood $[\theta - x, \theta + x]$ of the true parameter, it is at least as good with regard to a broad class of loss functions which includes absolute mean deviation ($L(x,\theta) = |x-\theta|$), variance ($L(x,\theta) = |x-\theta|^2$), and any other symmetric loss function where the loss is at least as large when x is farther away from θ than when it is closer to θ. It includes also the continuous entropy. Indeed, the latter is defined as: $H(X) = -\int_X f(x)Ln(f(x))$, and the majorization order implies the entropy order [19]. Therefore, for continuous unimodal distribution, the specificity order implies the entropy order.

Thus, the preference ordering of estimators based on greater possibility specificity around the parameter is indeed a strong preference ordering.

2.4 Relationships with other notions

The specificity order (defined by the fuzzy subset inclusion) is also related to second order stochastic dominance (SOSD) in connection with risk-aversion in decision theory. In this context, the notion of dispersion is translated to a notion of risk. Given two symmetric distributions with the same expected value, *G* is riskier than *F* if every risk-averse individual prefers *F* to *G*. One mathematical definition is:

$$F \le_{SOSD} G \Leftrightarrow \int_{-\infty}^t F(x)dx \le \int_{-\infty}^t G(x)dx, \forall t$$

We have: $\pi_X^M(x) \le \pi_Y^M(x) \Rightarrow X \le_{SOSD} Y$.

Proof: $\int_{-\infty}^t F(x)dx \le \int_{-\infty}^t G(x)dx$ is equivalent to:

$$U(t) = \int_{-\infty}^t (F(x) - G(x))dx \le 0, \forall t$$

We have: $U'(t) = F(t) - G(t) \le 0, \forall t \le 0$

due to $\pi_X^M(x) \le \pi_Y^M(x)$. Thus U is decreasing, and since $U(-\infty) = 0$, then $\forall t \le 0, U(t) \le 0$.

By a symmetric reasoning, $\forall t \ge 0, U(t) \le 0$, which completes the proof.

A link exists also with the Value at Risk notion $VaR_\alpha(X)$ that is defined by the 1-α quantile of X, i.e.:

$$P(X > VaR_\alpha(X)) = 1 - \alpha .$$

It is easy to see that:

$$\pi_X^M(x) \le \pi_Y^M(x) \Rightarrow VaR_\alpha(X) \ge VaR_\alpha(Y), \forall \alpha \in [0,1]$$

The specificity order is also related to notions of social science, i.e. the Lorenz curve and the Gini index defined by:

$$L_X(t) = \frac{1}{E(X)} \int_O^t F(x)dx \text{ and } G(X) = 1 - 2\int_O^1 L_X(t)dt$$

In fact, Lorenz dominance order is equivalent to second order stochastic dominance for distributions with equals means [19]. Thus, we have:

$\pi_X^M(x) \le \pi_Y^M(x) \Rightarrow L_X(x) \le L_Y(x), \forall x$ and,

$\pi_X^M(x) \le \pi_Y^M(x) \Rightarrow G(X) \ge G(Y)$.

The Gini index is a measure of statistical dispersion most prominently used as a measure of inequality of wealth distribution. A low Gini index indicates more equal wealth distribution, while a high Gini index indicates more unequal distribution.

3 Mean and median estimators

Let us consider the case where the parameter to be estimated is a location parameter of a continuous symmetric distribution function. This situation is often encountered in physical entity measurements where the variability of observations is due to variability of influence quantities or phenomena [3].

We considered hereafter the parameter estimation by using the mean and the median estimators.

3.1 Mean estimator

Suppose $X_1, X_2, ..., X_n$ are an *iid* sample from a distribution F. The mean estimator is defined: $\hat{\theta}_n = \left(\sum_{i=1}^n X_i\right) / n$.

Generally, the Gaussian distribution with a deviation σ is used for F , because, in this case the distribution of the mean estimator is also Gaussian but with a standard deviation $\sigma / \sqrt{n}$. Hereafter are plotted the possibility distributions associated to such mean estimators for $n = 1,3,30$ and $\sigma = 1$. As expected, the larger the sample size the more specific the possibility distribution.

ISBN: 978-989-95079-6-8

Let us now consider that the sample data are issued from a Cauchy distribution with a scaling parameter of *1*: $f(x) = \frac{1}{\pi(1+x^2)}$.

This distribution has no finite moments, but the median exists and can be used as a measure of central tendency. According to equation (1), the equivalent possibility distribution is defined by:

$$\forall x \le 0, \pi(x) = 1 + \frac{2}{\pi} arctg(x)$$

$$\forall x \ge 0, \pi(x) = 1 - \frac{2}{\pi} arctg(x)$$

The probability distribution of the mean estimator is also a Cauchy distribution with a scale parameter of *1* [1]. Thus the possibility distribution of the mean estimator of any number of data is the same as the one obtained for one datum. Therefore the specificity is not improved by increasing the sample size. Note that this is not in contradiction with the central limit theorem though the latter considers only distribution with finite variance (which is not the case of the Cauchy distribution).

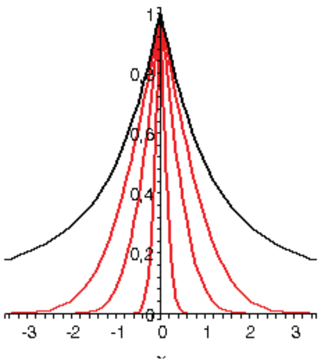

Figure 1: Possibility mean estimators for Normal (red) and Cauchy (black) probability distributions (*n=1, 3, 30*).

3.2 Median estimator

Let M_n be the median of $X_1, X_2, ..., X_n$, θ being the median of the distribution F from which the sample data are issued. For n odd, we have the following relationship [20]:

$$P(M_n - \theta \ge x) = \sum_{i=0}^{(n-1)/2} C_n^i F(x)^i (1-F(x))^{n-i}.$$

Therefore the expression of the possibility distribution of median estimator is:

$$\forall x \le \theta, \pi_{T_\theta(X)}(x) = 2(1 - \sum_{i=0}^{(n-1)/2} C_n^i F(x)^i (1-F(x))^{n-i})$$

$$\forall x \ge \theta, \pi_{T_\theta(X)}(x) = 2 \sum_{i=0}^{(n-1)/2} C_n^i F(x)^i (1-F(x))^{n-i}$$

The possibility median estimators for a Gaussian and a Cauchy distribution are plotted in figure 2 for $n = 1, 3$.

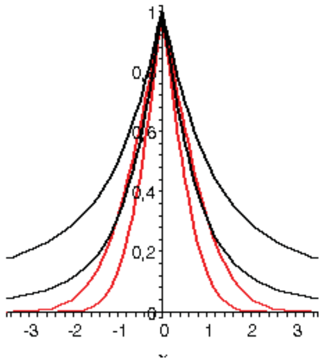

Figure 2: Possibility median estimators for Cauchy (black) and Normal (red) distributions (n=1, 3).

Note that the specificity increases with the sample size in both cases, which is remarkable for the Cauchy distribution.

3.3 Comparison

The preceding results show that the specificity of the possibility distribution associated to a median estimator increases with the sample size (with odd data) even for a long tail probability distribution such as the Cauchy distribution. It is not the case for the mean estimators. This seems indicate that the median estimator is better. Hereafter are plotted the possibility mean and median estimators for a few observations from a normal distribution.

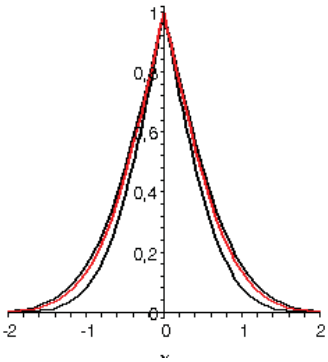

Figure 3: Possibility mean (black, n=2, 3) and median (red, n=3) estimators for a Normal distribution.

The possibility distribution of the median estimators with 3 observations is more specific than the possibility distribution of the mean estimator for 2 observations but less specific than the possibility distribution of the mean estimator for 3 observations. It seems that in general the possibility median estimator for *2n+1* data is more specific than the mean estimator for *2n* data.

ISBN: 978-989-95079-6-8

4 Conclusions

The paper has presented new links between continuous symmetric probability and possibility distributions. The interest of the possibility approach in the context of parameter estimation has been highlighted through the powerful concept of specificity ordering of possibility distribution. The latter allows sounded and intuitive comparisons between different estimators (e.g. mean and median). Further developments should consider dissymmetric distributions and other estimators.

References

[1] J. Rice, *Mathematical Statistics and Data Analysis*, Duxbury Press, 1995, 597 pages.

[2] A. Birnbaum, "A unified theory of estimation", *The Annals of Mathematical Statistics*, Vol. 32, No 1, 1961, p. 112-135.

[3] *Guide for the expression of uncertainty in measurement*, ISO 1993, 1993.

[4] J.W. Tuckey, "Standard confidence points", *Memorandum report 26,* Princeton University, 1949.

[5] D.R. Cox, "Some problems connected with statistical inference", *The Annals of Mathematical Statistics*, Vol. 29, 1958, p. 357-372.

[6] Z.W. Birnbaum, "On random variables with comparable peakedness", *The Annals of Mathematical Statistics*, Vol. 19, 1948, p. 76-81.

[7] G. Mauris, V. Lasserre and L. Foulloy, "A fuzzy approach for the expression of uncertainty in measurement", *Int. Journal of Measurement*, Vol. 29, No 3, March 2001, pp. 165-177.

[8] D. Dubois, L. Foulloy, G. Mauris and H. Prade, "Probability-possibility transformations, triangular fuzzy sets and probabilistic inequalities", *Reliable Computing*, Vol. 10, No 4, 2004, pp. 273-297.

[9] D. Dubois, H. Prade and S. Sandri, "On possibility/probability transformations", in*: Fuzzy Logic*, R. Lowen, M. Roubens, Eds, 1993, pp. 103-112.

[10] C. Baudrit and D. Dubois, “Practical representation of incomplete probabilistic information”, *Computational Statistics & Data Analysis*, Vol. 51, Issue 1, 2006, pp. 86-108.

[11] L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, Vol.1, No1, 1978, pp. 3-28.

[12] D. Dubois and H. Prade," On several representations of an uncertain body of evidence", in: *Fuzzy Information and Decision Processes* (M.M. Gupta, E. Sanchez, eds.), North-Holland, 1982, pp. 167-181.

[13] M. Kendall and A. Stuart, *The advanced theory of statistics*, Ed. Griffin and Co., 1977.

[14] P. Lévy, *Théorie de l’addition des variables aléatoires,* Gauthier-Villars, Paris, 1954, 416 pages.

[15] R.J. Hickey, "Continuous majorization and randomness", *Journal of applied probability*, No. 23, 1984, pp. 924-929.

[16] J.M. Fernandez-Ponce and A. Suarez-Llorens, "An aging concept based on majorization", *Probability in the engineering and informational science*, No 17, 2003, pp. 107-117.

[17] D. Dubois and H. Hüllermeier, "Comparing probability measures using possibility theory: a notion of relative peakedness", *Int. journal of Approximate reasoning,* No 45, 2007, pp. 364-385.

[18] J.T. Hwang, "Universal domination and stochastic domination: estimation simultaneously under a broad class of loss functions", *Annals of Statistics,* No 13, 1985, pp. 295-314.

[19] A. Müller and D. Stoyan, *Comparison methods for stochastic models and risks*, Wiley, 2002, 330 pages.

[20] H.A. David, *Order Statistics*, Wiley, 2003.

ISBN: 978-989-95079-6-8