STATISTICAL RISK AND THE MANAGEMENT DECISION

Irina-Maria Dragan

Abstract

The volatility becomes a constantly characteristic of the present economic environment. Considering those circumstances, the decisions are affected by risks in an expanding capacity. Since the management decisions are typically taken on the basis of incomplete information, based on some partial data and in conditions of uncertainty, these are associated with errors. Because the risk is regarded as a probability, therefore it is worth to investigate the nature of what is called the statistical risk. It plays an important role in the framework of statistical inference. The statistical feature of the decisional area determines that the error's probabilistic measurement to be made by the risks. In this paper there are depicted and adjusted the risks, such as Type I, Type II and Type III. The variance resulting from measurement error becomes noise and thus it could decrease the power level. To be more specific, since a high degree of measurement error hinders the condition of the variable from being correctly indicated, it drags down the possibility of correctly detecting the effect under study. Those risk types are affecting the alternative hypotheses of the decision making processes and also the Taguchi risk and the potential index of a process.

Key words: statistical risk, risk of errors, decisional process

JEL Code: C44, D81

Introduction

One of the simplest decisional problems under *uncertainty conditions* is to accept or to reject a statistical hypothesis, hypothesis that can be true or false. Uncertainty in the decisional process occurs due to sampling: we work with a part (or with parts of the origin population), consequently, with samples of a certain "whole" (set, batch, population etc.) and not with the whole collectivity, the decision to accept/reject being made on the basis of examining only that part represented by the *sample*. The error is the wrong decision and the probabilistic measure of making the error is the risk. In this paper was developed the subject of uncertainty of decision and of risk in general, and particularly, of the statistical one.

1 Uncertainty and Management Decision

The usual definition of uncertainty says that we are dealing with something uncertain, undetermined and doubtful: uncertainty is a *doubt* about something or somebody. Certainty is considered as the opposite of uncertainty and signifies the fact that a certain fact, event, situation will take place or is taking place under our sights. We mention among the uncertainty sources: incomplete, partial information on a certain entity, lack of information, inadequate rendition of information, wrong assignment of causality.

Malita and Zidaroiu have published a work where they present the mathematical formalization of the decisional process and of the decision – making process (in the economic, medical, technological, administrative etc. field) having as theoretical fundamentals – among others – the probabilistic – statistical procedure and that one related to the utility theory. If we are interested of specific fields, as, for instance, the metrology: we are dealing with the *measuring* uncertainty, that, according to "International Vocabulary of Basic and General Terms in Metrology" 1992 is a *parameter* associated to the result of a measuring which characterize the dispersion of values (this parameter can be a standard deviation, a certain interval de statistical coverage). Here, uncertainty gets a correct contour by a statistical indicator, by formulae, expressing a certain fact – namely: how far or how close we are to the *real value* of the physical magnitude subjected to the measuring process (Hopkin, 2012).

We do not have, in mathematical statistics, a definition of the idea of uncertainty of a certain system: entropy, negentropy, information, informational energy etc. provide an image of the respective system state as against an eventual "standard system" and obviously in relation to the known subject, to the anthrop-social element. Such, uncertainty appears as the fruit of human ignorance, and its manifestation form is the variability, which if it is not controlled, can generate what we usually call a *risk*: namely, to make a wrong decision into a given situation, where the required information is distorted because of this exaggerated variability. Barsan-Pipu and Popescu in a profile work, consider that *risk* is however an *uncertain event*, which can very well not to occur. Uncertainty is of two types: ontological and cognitive one. The ontological uncertainty (or the nondeterministic one) deals with problem, abandoning the idea that the future states come from current states. The calculation using such probabilistic knowledge will lead to choosing a solution, which might not be the best solution, but, for sure, the choice was made because the solution has the highest probability to be the best solution. Anyway, the selected solution could be a worse solution, but it was a right, rational selection, with the highest probability to be the best solution. The probabilistic thinking is the way of being rational and of making the right decisions into a partially nondeterministic world. The cognitive uncertainty expresses the limits of our knowledge, thus, the decision – maker assigns manageability to different events. But, unlike the uncertainty, the risk is characterized by the possibility to be quantified by probability laws. Consequently, a discussion about the concept of risk is required.

2 Uncertainty and the connection with the decision under risk

The risk is a phenomenon which affects and influences any human action. This can be defined as a commitment bearing an uncertainty, due to the probability of money – making or money - losing. The social, economic and cultural progress would not have been possible without assuming certain risks in decision – making, both at personal level and within organizations. A rigorous definition of risk says that "the context where an event occurs with a certain probability or where the extension of the event follows a probability distribution"[1]. This definition outlines that the actions taken at a given moment by a person or by an organization, lead to an event or to a set of events, which a probability or a distribution of occurrence probabilities are associated to and the risk is nothing else but what results from the combination of the occurrence probabilities and their consequences. Etymologically, the word comes from the Latin *re-secare* - which means a balance breaking[2] (in French: risqué). Further on, we present some definitions (dictionary) related to this rendition.

Mic Dictionar Enciclopedic (Stiintifica and Enciclopedica Publishing, Bucuresti, page. 809): "Risk: danger, possible inconvenience". Then, it follows: the *"contractual risk*", i.e.: the debtor must bear the injury consequences of releasing the creditor from the obligations he had to him, as a result of non-performing, by debtor, of his duties from causes which are not ascribed to him. Here, we meet also the notion of *working risk* – the owner shall bear the damages resulting from the loss or from the destruction of work by force majeure. We shall also mention the notion of *insured risk*. An immediate derivative is *risky*, i.e.: full of risks, dangers exposed, unsafe etc.

Merriam – Webster's Collegiate Dictionary (10th Edition 1994, Springfield, Mass., pp. 1011) providing historical details, i.e.: when "risk" term (and its derivatives) started to be used in English: the possibility that something bad or unpleasant (such as an injury or a loss) will happen; someone or something that may cause something bad or unpleasant to happen –

[1] Pierce, D. (edit) *Macmillan Dictionary of Modern Economy*, Codecs Publishing, page 353, 2000
[2] http://www.oed.com/

see also Security Risk; a person or a think that someone judges to be a good or bad choice, for instance a loan etc. Larousse de Base (Dict. d'apprentisage du français), Libr. Larousse Paris, page 723): "Risk" is a *danger* that can be foreseen; an (abstract) *disadvantage* we are exposed to; an action involving risks is called *risky* or hazardous (hasardeuse).

Terry Lopez states may be the most concise but still comprehensive definition: "risk means danger to suffer a loss". The change of this image comes firstly from the introduction of the uncertainty concept into the rational decision model. Th. Craiovenu remarks the fact that "when uncertainty reaches high levels, it must be diminished, but the most undesirable level of uncertainty in life is that of full certainty.

Some conclusions appear clearly from all these definitions: the risk association to human factor ("the cars do not have troubles" and the lions of African savannahs "do not suffer damages" but "misfires" in their chase after impalas); the risk is indissolubly related to uncertainty or loss. The risk can or can not occur itself through the so-called *risk factors*, which, when they effectively operate, can cause loses of diverse nature and magnitudes. We have also to mention that despite the fact that there is a risk, in a certain situation, this is possible not to occur (the decision-maker is acting "on chance" and he is fortunate enough not to suffer losses, damages etc.). The authors' opinion is that risk is present in all human activities and for this reason the risk theory is a wide and difficult field of research. The mathematical modeling of risk has as start point the assumption that risk can be assimilated to the possibility of suffering a loss. As this possibility can be quantitatively expressed by *probability*, risk appears as a probability occurrence function of an undesired phenomenon, and also the presence of the adverse effects of this event.

3 Statistical Measurement of Risk

Statistics – according to its numerous definitions, covers this too: "the science of decision-making under uncertainty (Lovric, 2011; Lindley, 2000). Statistical hypotheses are statements on one or several parameters, or on some distributions which are going to be validated by statistical tests. The decision on null hypothesis is made on the basis of statistical test algorithms. The statistical test being itself built with random elements, the decision-making bears an error risk. The null hypothesis (H_0) refers to statements to be subjected to testing, while the alternative hypothesis (H_1) refers to statements which will be accepted if the null hypothesis is rejected. Testing the hypothesis that the mean μ of a random variable X of a population is not lower than a given value μ_0: $H_0(\mu \geq \mu_0)$ and $H_1(\mu < \mu_0)$. Testing the hypothesis that the proportions of unconformable individuals of two populations p_1 and p_2 have the same (unspecified) value: $H_0(p_1 = p_2)$ and $H_1(p_1 \neq p_2)$. Testing the hypothesis that a random variable *X* has a normal distribution (with unspecified parameters). The alternative hypothesis: the distribution is not normal.

The critical region R_c is the region R_c of rejecting the statistical hypothesis and is called the critical region of the Euclidean space R^n. In order to build the critical region R_c of a test, there are usually used certain sample indicators of the form $E(x_1, \ldots, x_n)$ and the probability α to make a type I error. The critical region is defined by set R_c that satisfies the condition as the type I error probability $E \in R_c$, when hypothesis *h* is true, to be equal to α, i.e.: $P[E(x_1, x_2, ..., x_n) \in R_c; H] = \alpha$. Type II error probability is defined by relation: $P[E(x_1, x_2, ..., x_n) \in R_c^* \text{ non } H] = \beta$ where "non H" is the negation of hypothesis H. The critical regions are determined so that, if the null hypothesis is the true probability of rejecting the null hypothesis, it must be at most equal to a given value α, generally a low one (for example: 5% or 1%). **Type I error probability.** The value of this probability is always lower or equal to the test significance level. In certain situation, it is also called: "type I risk". **Type II error probability**. Its value depends on the real situation and can be calculated only if the alternative hypothesis is enough specified. Sometimes, it is also called: "type II risk". **Power of a test**. This probability, generally denoted by $\Pi_t = (1 - \beta)$, corresponds to the null hypothesis, when it is false. We can also recall the fact present in literature, on a proposal, who introduces also a Type III error, that, according to his opinion, consists in solving *false problems*. We consider that a better clarification is required: what is, in fact, a "false" problem? Is the problem stated incorrectly or simply without any connection with the decision – maker's desire? The author does not clarify this aspect.

4 Some Models of Decision - Making Risk in Processes Management

From a statistical point of view, uncertainty and implicitly risk are modeled with the aid of some random variables. In this way, risk is considered a random (continuous) variable $X : \Omega \rightarrow R$ where (Ω, K, P) is a probability area. If $V(\Omega)$ is the set of all variables defined on Ω, and $M \subset V(\Omega)$ is a subset of $V(\Omega)$, then, a function $r : M \rightarrow R$ is a so-called measure $r(X)$ of risk X. Some examples related to the meaning of X: net profit of an investment, the loss suffered by an employee into a certain conjuncture, exceeding the alarm quota of a pollutant concentration etc. (Vose, 2008; Voda, 2009). Consequently, if X is the investment profit, the associated risk is:

$$r_1(X) = Var(X) = E\left[(X - m)^2\right] \quad (1)$$

where $m = E(X)$ is the mean value of variable X. At the same time,

$$r_2(X) = E\left[(T - X)_+^2\right] \quad (2)$$

where T is the a so-called "failure limit" (a name given by Roy[3]), the investor wishes to have always a profit above the threshold T). If X is rendered as getting beyond the pollutant concentration with respect to a certain maximum level T, then, the polluting risk is:

$$r_3(X) = E\left[(X - T)_+^p\right],\ p \in \{1,2,...\} \quad (3)$$

The used notation derives from "*Lower Partial Moment*" indicator of α order with respect to the limit *T*:

$$LPM_\alpha(T, X) = E\left[(T - X)_+^\alpha\right] = \int_{-\infty}^{T} (T - x)^p \, dF_X(x) \quad (4)$$

where $F_X(x)$ is a distribution function of *X*. If *T* is even $m = E(X)$, and $\alpha = 2$, then, the so-called semi variance is obtained:

$$SV(X) = E\left[(m - X)_+^2\right] \quad (5)$$

As can be noticed, we have to do with risk models using dispersion $(r_1(X))$ or dispersion with respect to a given value *T*. Further on, we shall present some variants of modeling the risk in processes management, mainly, for monitoring their capacity (Wright, 1995).

Taguchi risk. In fact, the forms $r_2(X)$ and $r_3(X)$, (relations 2, 3) are nothing else but the *average risks* associated to a quadratic form $L(x,T) = k(x - T)^2$ used by C.F. Gauss since 1809 (Kackar, 1985) and revitalized by Genichi Taguchi in "quality context": *T* is the target value of the measurable characteristic *X*, and *L*(*x*,*T*) is the loss quality function. Then:

$$R_T(x) = E[L(x,T)] = \int_D L(x,T) f(x) dx \quad (6)$$

is the Taguchi risk, i.e.: the mean value of variable *L*(*x*,*T*), where $f(x) = F_X^{'}(x)$ and *D* is the definition field of *X* (usually $[0,+\infty)$). Genichi Taguchi revitalized Gauss quadratic function $f_{(x)} = a(x - x_0)^2,\quad a > 0,\ x \geq x_0 \geq 0$ and associated it to the so-called quality loss:

$$\mathrm{L}(x_0;\mathrm{T}) = k(x_0;\mathrm{T})^2,\quad k > 0,\ x_0, \mathrm{T} \in \mathrm{R} \quad (7)$$

[3] Roy, A. "Safety First and the Holding of Assets", *Econometrica*, 20(3): 431-449, 1952

where x_0 is the measured value of the quality characteristic (X) and (T) is its target value (k is a constant depending on the specific case at hand).

If $f(x;\theta)$ is the density of X $(x \in \mathrm{D} \leq \mathrm{R}, \quad \theta \in \mathrm{R})$ then the average value

$$\mathrm{E}[\mathrm{L}(x;\mathrm{T})] = \int_{\mathrm{D}} \mathrm{L}(x;\mathrm{T}) f(x;\theta) dx \qquad (8)$$

is called Taguchi type risk (Kackar, 1985). Taking into account (7) we may write (8) as

$$\mathrm{E}[\mathrm{L}(x;\mathrm{T})] = k\left[\mathrm{Var}(\mathrm{x}) + (\mathrm{E}(\mathrm{x}) - \mathrm{T})^2\right] \qquad (9)$$

and if X is normally distributed $\mathrm{N}(\mu;\delta^2)$, we have:

$$\mathrm{E}[\mathrm{L}(x;\mathrm{T})] = k\left[\delta^2 + (\mu - \mathrm{T})^2\right] \qquad (10)$$

The empirical risk (denoted by $\hat{\mathrm{R}}_{\mathrm{T}}(x)$) is therefore:

$$\hat{\mathrm{R}}_{\mathrm{T}}(x) = k\left[\mathrm{s}^2 + (\bar{x} - \mathrm{T})^2\right] \qquad (11)$$

where $\bar{x}$ and s are the well-known sample statistics. There is a straight forward link between Taguchi's risk and his own process capability index $\hat{\mathrm{C}}_{\mathrm{pm}}$ (Chan et al, 1988):

$$\hat{\mathrm{C}}_{\mathrm{pm}} = \frac{\mathrm{USL} - \mathrm{LSL}}{6 \cdot \sqrt{\mathrm{s}^2 + (\bar{x} - \mathrm{T})^2}} \qquad (12)$$

where USL = Upper Specified Limit and LSL = Lower Specified Limit of the given quality characteristic $\mathrm{X} \sim \mathrm{N}(\mu, \delta^2)$ with T as its target value. We may write hence immediately:

$$\hat{\mathrm{R}}_{\mathrm{T}}(x) = k\left(\frac{\mathrm{USL} - \mathrm{LSL}}{6} \cdot \frac{1}{\hat{\mathrm{C}}_{\mathrm{pm}}}\right)^2 \qquad (13)$$

If USL - LSL = 6s - that is the minimal level for admissible process capability, we get:

$$\hat{\mathrm{R}}_{\mathrm{T}}(x;\mathrm{T}) = k\left(\frac{\mathrm{s}}{\hat{\mathrm{C}}_{\mathrm{pm}}}\right)^2 \qquad (14)$$

and we draw the conclusion that the Taguchi's risk can be regarded as a function of the length of the specified interval USL - LSL measured in standard deviations units.

The theoretical Taguchi risk corresponding to (14) is:

$$\hat{\mathrm{R}}_{\mathrm{T}}(x) = \mathrm{k}\left(\frac{\delta}{\hat{\mathrm{C}}_{\mathrm{pm}}}\right)^2 = \frac{\mathrm{k}\delta^2}{\hat{\mathrm{C}}_{\mathrm{pm}}{}^2} \qquad (15)$$

Denoting $k\delta^2 = M$ and $C_{pm} = X$, we shall have a hyperbolic dependence of the type $R = M/X^2$. If in C_{pm}, the true mean-value μ is just the target T, then C_{pm} becomes the classical potential index of a process, namely $C_{pm} = (USL - LSL)/6\delta$.

Wright risk. Wright (1995) introduced an index which takes into account the asymmetry of the distribution:

$$C_s = \frac{d - |\mu - T|}{3\sqrt{\sigma^2 + (\mu - T)^2 + |\mu_3 / \sigma|}} \tag{16}$$

where T = (LSL + USL)/2, d = (USL - LSL)/2 and $\mu_3 = E(x-\mu)^3$ - the third central moment which describes the departure from symmetry.

Luceno risk. Alberto Luceno proposed the index:

$$C_{pc} = \frac{USL - LSL}{6\sqrt{\pi/2} \cdot E|X - m|} \tag{17}$$

where m = (LSL + USL)/2 - this value may be (or may be not) a target value for X.

Crisan risk. D.M. Crişan uses the Kurtosis to correct the potential index, namely:

$$\hat{C} = \frac{USL - LSL}{6s \cdot \sqrt{1 + |3 - b_2|}} \tag{18}$$

where $b_2 = \frac{1}{n}\sum (x_i - \bar{x})^4 / s^4$ is the sample coefficient of Kurtosis. We have to notice that the new index takes into account the curve type (platykurtic or leptokurtic as $b_2 < 3$, respectively, $b_2 > 3$, namely with a flatter peak around its mean, which causes thin tails within the distribution, respectively with a higher peak around the mean compared to normal distributions, which leads to thick tails on both sides). Consequently:

$$\hat{C} = \begin{cases} \dfrac{T_s - T_i}{6s \cdot \sqrt{4 - b_2}}, b_2 < 3(platykurtic) \\ \dfrac{T_s - T_i}{6s \cdot \sqrt{b_2 - 2}}, b_2 > 3(leptokurtic) \end{cases} \tag{19}$$

In case of normal law, we have $b_2 = 3$ and consequently, the Crisan's index becomes again the classical one $\hat{C}_p$. Barsan-Pipu and Popescu outline an interesting relationship between the Taguchi average loss and the Taguchi potential index, assuming a certain statistical performance level of the NPS process:

$$E[L(X,T)] = K \cdot \left(\frac{NPS \cdot \sigma}{C_{pm}} \right)^2 \tag{20}$$

Concrete examples of calculations are given in the cited work – calculating inclusively the costs of such an approach (see pages 107 – 111). ISO document ISO/DIS 3534-2/2004 (ϕ 1.2 "Statistical process management", pages 11 - 30) introduced the so-called reference interval $x_{99.865\%}$ - $x_{0.135\%}$, where $x_{a\%}$ is the a% - fractile of the distribution. This interval is then used to construct "process performance index"

$$P_p = \frac{USL - LSL}{x_{99.865\%} - x_{0.135\%}} \qquad (21)$$

which should be used when the process is not in statistical control (for normal distribution, the reference interval is 6σ or 6s). Analogously, one may introduce

$$P_{pk_L} = \frac{x_{50\%} - LSL}{x_{50\%} - x_{0.135\%}} \text{ and } P_{pk_U} = \frac{x_{50\%} - LSL}{x_{50\%} - x_{0.135\%}} \qquad (22)$$

and finally $P_{pk} = \min\{P_{pk_L}, P_{pk_U}\}$

Caulcutt risk. Caulcutt (1995) uses the index (page 155):

$$C_{pu} = \frac{USL - Median}{3s - Median} \qquad (23)$$

in order to describe a Skewed distribution.

Vännman risk. It is important to notice that Kerstin Vännman unified some of these indices:

$$C_p(u, v) = \frac{d - u|\mu - M|}{3\sqrt{\sigma^2 + v(\mu - T)^2}} \qquad (24)$$

where *u, v* are real numbers, *d* = (USL - LSL)/2, *M* = (LSL + USL)/2, *T* - target value (which may be *M* sometimes). Hence, we get easily: $C_p(0, 0) = C_p$ (classical potential index); $C_p(1, 0) = C_{pk}$ (Kane); $C_p(0, 1) = C_{pm}$ (Taguchi); $C_p(1, 1) = C_{pmk}$ (Peern - Kotz - Jonhson).

Conclusion

The management process is strongly linked to the micro and macroeconomic environment, full of uncertainties and the decisions are affected by risks. Mainly, these are occurring especially in decision-making, and as they affect only one of the parts, they generate different risks, statistically measured by β – type II risk.

These risks arise and are of importance in the company relations, with respect to external environment, decision-making on contracts deliveries, purchases, economic exchanges of goods and services etc. As for the management internal process, the control of processes (i.e.: the manufacturing processes in a factory, the bureaucratic processes in a bank or in administrative institutions, the flow processes in a hospital etc.) are subjected to a certain risk, probabilistically measured, to get out of normal limits. This issue is measured by its capacity and exceeding the capability limits is measured by different risks as Taguchi, Waright, only to cite the most known ones. According to the process state, a whole literature was developed, that provides solutions to the diversity of practical situations which are of real help in the operational management of processes.

References

Borgonovo, E., "A Methodology for Determining Interactions in Probabilistic Safety Assessment Models by Varying One Parameter at a Time", *Risk Analysis*, 30(3):385-99, 2010

Caulcutt, R. *Achieving Quality Improvement. A Practical Guide*, Chepman and Hall, London, 1995

Chan, L. K., Cheng, S.W. and Spiring, F. A. "A New Measure of Process Capability: Cpm", *Journal of Quality Technology*, 20 (3): 162-175, 1988

Christoph H. L., Michael E. S. and Elaine M. B., „Diagnosing Unforeseeable Uncertainty in a New Venture", *Journal of Product Innovation Management*, 25(1): 28–46, 2008

Crowder, M., "Characterizations of Competing Risks in Terms of Independent-Risks Proxy Models", *Scandinavian Journal of Statistics,* 27(1): 57–64, 2000

Hopkin, P., *Fundamentals of Risk Management*, 2nd ed., Kogan-Page, 2012

Isaic-Maniu, I. *Caracterizarea statistica a riscului*, ASE Publishing, 2006

Kackar, R. N. "Off-line Quality Control, Parameter Design and the Taguchi Method", *Journal of Quality Technology*, 17(4): 176-188, 1985

Lindley, D. "The Philosophy of Statistics", *Journal of the Royal Statistical Society,* Series D 49 (3): 293-337, 2000

Lovric, M. (edit.) *International Encyclopedia of Statistical Science*, Springer, 2011

Vose, D. *Risk Analysis: A Quantitative Guide*, John Wiley & Sons, London, 2008

Voda, V. "Some Comments on Statistical Risks", *Reliability &Risk Analysis: Theory & Applications*, 1(12) vol.2: 121-127, USA, 2009

Wright, A. P. "A Process Capability Index Sensitive to Skewness", *Journal of Statistical Computation and Simulation*, 52: 195-203, 1995

Contact

Irina-Maria Dragan
Academy of Economic Studies Bucharest
Calea Dorobantilor nr. 15-17, sector 1, Bucuresti, cod 010552
irina.dragan@csie.ase.ro