Discrete Variable's Measurement Using Non-numerical, Inexact, and Incomplete Information on Probabilities' Distribution

N.V. Hovanov
St.Petersburg State University
nick@polyidea.com

When measuring a discrete variable x, whose possible values constitute a finite set $X = \{x_1,...,x_n\}$, one uses, as a rule, the stochastic model, suggesting that the observed results obtained in the measurement of the variable are realizations of a random variable $\tilde{x}$, this variable being specified by a probabilities' distribution $p = (p_1,...,p_n)$, where p_i is a probability that the corresponding value x_i will appear. Within this model's framework one usually chooses a mathematical expectation $\mu = E\tilde{x} = x_1 \cdot p_1 + ... + x_n \cdot p_n$ of the random variable $\tilde{x}$ as an estimation of the discrete variable under measuring.

Unfortunately, in a real measurement practice, we usually have to deal with an *uncertainty* of probabilities' vector $p = (p_1,...,p_n)$ determination, this uncertainty being manifested in the fact that the vector is represented within an accuracy of a certain set $P(n;I)$ of all admissible (from the point of view of an additional information I) probabilities' vectors. In so doing, the information I, which determines the set $P(n;I)$, is often *non-numerical, inexact, and incomplete* one. For modeling of the uncertainty of a probabilities' vector specification we propose to use the Bayesian strategy of "*randomization of uncertainty*". In the case considered, the uncertainty of the choice of a vector $p = (p_1,...,p_n)$ from the set $P(n;I)$ is modeled by a random vector $\tilde{p} = (\tilde{p}_1,...,\tilde{p}_n)$, which takes values from the set $P(n;I)$. The obtained *randomized (random) probabilities* $\tilde{p}_1,...,\tilde{p}_n$ can be substituted into the expression for the expectation μ, what will give the *randomized mathematical expectation* $\tilde{\mu} = x_1 \cdot \tilde{p}_1 + ... + x_n \cdot \tilde{p}_n$, this randomized expectation being used as a model for the uncertainty of the discrete variable's estimation.

The report gives explicit formulae for mathematical expectation $\bar{\mu}(I)$ and variance $\bar{\sigma}^2(I)$ of the randomized expectation $\tilde{\mu}$, these formulae being used for a *numerical image* $(\bar{\mu}(I),\bar{\sigma}^2(I))$ *of non-numerical, inexact and incomplete information* I, which is possessed by an investigator, who is carrying out a measurement of a discrete variable.