Methodology of predictions and their quality

A. L. Semenov, An. A. Muchnik

Council of Cybernetics at Computing Center
of Russian Academy of Sciences

We consider the following interpretation of scientific explanation and theory building. The real world (that can be a part of the whole reality) is a tree of states with some measure (probability distribution) on it. A hypothesis gives us a measure on the same tree also.

Let us have a family of hypotheses (finite, infinite, or even continuous set) including one corresponding to the real world (that equals to its measure). By methodology we mean a method that given any such family (with no indication to the real world) produces a hypothesis. We say (informally) that methodology is good, if the hypotheses produced by it are close to reality.

Every hypothesis μ for any node x sets probability distribution on immediate followers of this node:

$$p(a) = \mu(xa|x) = \mu(xa)/\mu(x).$$

Our characteristics of closeness of probability distributions for one node is based on Hellinger distance:

$$\rho(p,q) \;=\; \sum\nolimits_a (\sqrt{p(a)} - \sqrt{q(a)})^2 \;=\; 2 - 2\sum\nolimits_a \sqrt{p(a)q(a)}.$$

Characteristic of closeness of hypothesis ν and μ along the path y in the tree is

$$r(\mu,\nu,y) = \sum\nolimits_k \rho(\,\mu(\cdot|y_{1:k}),\;\nu(\cdot|y_{1:k})).$$

Global characteristic of closeness of hypothesis ν which is produced by methodology and hypothesis μ corresponding to the real world is

$$R_\alpha(\mu,\nu) = \;\mathbf{E}[\exp(\alpha r(\mu,\nu,y))],$$

where $\mathbf{E}$ — probabilistic average by measure μ, and α — an auxiliary parameter. The quality of methodology is measured by supremum of α, such that $R_\alpha(\mu,\nu)$ is finite in all cases.

Ray Solomonoff who originated this field of research proposed the methodology of (arithmetical) mean:

$$\nu = \frac{1}{k}\sum_{i=1}^{k} \mu_i.$$

Another methodology (best-buy) can be based on choosing the most probable move at any mode: $\nu(xa|x) = \mu_i(xa|x)$, where $\mu_i(x) = \max_j \mu_j(x)$.

We consider the case of two hypotheses only and prove that the mean methodology (maximum of α is equal to $\frac{1}{2}$), as well as best-buy methodology ($\sup\alpha \le \frac{1}{2}$) are not optimal. The optimal methodology is found (with $\sup\alpha = 1$).