Estimation of nonspecifity in the Dempster-Shafer theory.

Pavla Kunderová
Departement of mathematical analysis and numerical mathematics,
Palacký University Olomouc, Czech Republic

Nonspecifity is one of the measures of uncertainty in the Dempster-Shafer theory. Point and interval estimations of nonspecifity were suggested in this article.
Keywords: Dempster-Shafer theory, uncertainty, nonspecifity.

1. Terminology and notations.

Let X denote a universal set under consideration assumed here to be finite. Semantically X represents the set of all possible answers to a question (all possible states of a system, all possible diagnoses). We assume that elements of X are mutually exclusive answers (states, diagnoses). Let $\mathcal{P}(X)$ denote the power set of X. Subsets of X are identified with propositions ($A \subseteq X$ is identified with the proposition "the true answer is in A"). The Dempster-Shafer theory is based upon a function

$$m: \mathcal{P}(X) \longrightarrow \langle 0, 1 \rangle$$

such that

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{A \subseteq X} m(A) = 1 \quad .$$

This function is called *a basic belief assignment* (or just basic assignment). The value $m(A)$ represents the degree of belief that a specific element of X belongs to set A, but not to any particular subset of A. Every set $A \in \mathcal{P}(X)$ for which $m(A) \neq 0$ is called *a focal element*. The pair (F, m), where F denotes the set of all focal elements of m, is called *a body of evidence*.

Two distinct types of uncertainty are subsumed in the Dempster-Shafer theory. One of them is well characterized by the name *nonspecifity*. This type of uncertainty is properly measured by a function N defined by the formula

$$N(m) = \sum_{A \in F} m(A) \log_2 |A| ,$$

where $|A|$ denotes the cardinality of the focal element A. This function was proven to be unique under appropriate requirements [3].

It measures nonspecifity of a body of evidence in units that are called *bits*.

One bit of uncertainty expresses the total ignorance regarding the truth or falsity of one proposition. It is obvious that the function $N(m)$ attains its minimum, $N(m)=0$, if and only if all focal elements are singletons ($N(m) = 0$ for all probability measures). The maximum of $N(m)$ is attained when $m(X) = 1$, then $N(m) = \log_2 |X|$. Hence the range of $N(m)$ is

$$0 \le N(m) \le \log_2 |X| .$$

2. **Estimations of $N(m)$.**

There is a natural question: how to estimate the unknown value of $N(m)$ in the practice?

Let $Y = \{A_1, \ldots, A_r\}$ be a system of all non-empty subsets of the universal set X.

Let us repeate n independent experiments such that the result of each experiment consists in determination of one and only one element of a system Y.

Example.

A patient is suffering from a certain disease belonging to a set X. He will be examined by n (independent) doctors. Each of them determines one non-empty subset of X to which, according to him, the disease of the patient belongs.

Let ξ_{ji} be a random variable defined as follows:

$\xi_{ji} = 1$ in case the subset A_j is chosen in the i-th experiment,

$\xi_{ji} = 0$ in case the subset A_j is not chosen in the i-th experiment.

Let
$$Z_j = \sum_{i=1}^{n} \xi_{ji}, \qquad j=1,\dots,r,$$

thus the random vector $(Z_1,\dots,Z_r)$ has the multinomical distribution with parameters $n, p_1=m(A_1),\dots,p_r=m(A_r)$.

<u>*Theorem 1.*</u>

The random variable
$$\hat{m}(A_j) = \frac{Z_j}{n}, \qquad j=1,\dots,r,$$

is an unbiased estimation of $m(A_j)$, $\quad j=1,\dots,r,$

and the random variable

$$\hat{N}(m) = \sum_{A\subseteq X, A\neq\emptyset} \hat{m}(A)\log_2|A|,$$

is an unbiased estimation of a nonspecifity $N(m)$.

Proof. Obvious by using the definition of Z_j, $j=1,\dots,r$.

In the following we shall derive a confidence interval of the nonspecifity $N(m)$. Let us consider two probability distributions

$$P = \{p_1 = m(A_1),\dots,p_r = m(A_r)\}, \qquad \hat{P}=\{\hat{p}_1 = \hat{m}(A_1),\dots,\hat{p}_r = \hat{m}(A_r)\},$$

on the space
$$Y = (A_1,\dots,A_r) = (y_1,\dots,y_r).$$

Let us consider a random variable ξ defined on the probability space (Y,P) and its mathematical expectations under $\hat{P}$ and P

$$E_{\hat{P}}(\xi)=\sum_{i=1}^{r} \xi(y_i)\hat{p}_i, \qquad E_P(\xi)=\sum_{i=1}^{r} \xi(y_i)p_i.$$

and the variance under P

$$\sigma_P^2(\xi) = \sum_{i=1}^{r} [\xi(y_i)]^2 p_i - [E_P(\xi)]^2.$$

Lemma.

If $\sigma_P^2(\xi) > 0$ then $\sqrt{n}\,[E_{\hat{P}}(\xi) - E_P(\xi)]$ tends in law to $N(0,\sigma_P^2(\xi))$.

Proof. In [1],pg.13.

Let us define on (Y,P) a random variable ξ by

$$\xi(y_i) = \log_2|A_i|, \qquad i = 1,\dots,r,$$

then

$$E_{\hat{P}}(\xi) = \hat{N}(m), \qquad E_P(\xi) = N(m),$$

and

$$\sigma_P^2(\xi) = \sum_{i=1}^{r} [\log_2|A_i|]^2 m(A_i) - [N(m)]^2.$$

For the sake of simplicity we denote the variance $\sigma_P^2(\xi)$ by the symbol σ_m^2 in what follows.

Theorem 2.

If $\sigma_m^2 > 0$ then the random variable

$$\sqrt{n}\,[\ \hat{N}(m) - N(m)\ ]$$

tends in law to the normal distribution $N(\ 0,\sigma_m^2\ )$.

Proof. The statement follows by applying the lemma and the definition of the random variable ξ.

Corollary.

For sufficiently large number n and for $\sigma_m^2 > 0$ holds

$$P[\ \hat{N}(m) - u_{1-\alpha/2}\ \sigma_m [\ n\ ]^{-\frac{1}{2}} \leq N(m) \leq \hat{N}(m) + u_{1-\alpha/2}\ \sigma_m [\ n\ ]^{-\frac{1}{2}}\ ] = 1-\alpha,$$

where u_α denotes the α-quantile of the normal distribution $N(0,1)$.

Remark.

The unknown value of σ_m^2 could be replaced by its estimation

$$\hat{\sigma}_m^2 = \sum_{i=1}^{r} \hat{m}(A_i)[\log_2 |A_i|]^2 - [\ \hat{N}(m)\ ]^2.$$

References.

[1] J.Feistauerová,I.Vajda :Testing System Entropy and Prediction Error Probability.In:IEEE Trans.Syst.Man,Cybern.,1992.

[2] P.Hájek,D.Harmanec:On belief functions.In:Advances in AI,Lecture notes in AI,Vol 613,1992,pp.286 - 307.

[3] G.J.Klir,T.A.Folger:Fuzzy sets,uncertainty and information,Prentice Hall,Englewood Cliffs (N.J.),1988.

[4] G.J.Klir,A.Remer:Uncertainty in the Dempster-Shafer theory: a critical re-examination.Int.J.General Systems,Vol.18,pp.155-166.