ON DILATION

1. Good on the Principle of Evidence

Good (1967) provides a short mathematical argument to justify the value of acquiring new evidence in the context of rational decision making:

- If the cost of making new observations is ignorable, then it is alway beneficial to do so in the sense that the expected utility of the best action calculated with the information from new evidence is at least as great as that of the best action calculated without any further observations.

Table 1. Decision matrix.

	H_1	H_2	$\cdots$	H_n
A_1	$u_{1,1}$	$u_{1,2}$	$\cdots$	$u_{1,n}$
A_2	$u_{2,1}$	$u_{2,2}$	$\cdots$	$u_{2,n}$
$\vdots$			$\ddots$	
A_m	$u_{m,1}$	$u_{m,2}$	$\cdots$	$u_{m,n}$

Let $E_1, \dots, E_t$ be possible outcomes of, say, a new experiment T regarding $H_1, \dots, H_n$, and let $\Pr(\cdot)$ be the agent's prior.

$p_i = \Pr(H_i)$**:** the agent's prior over H_i.
$p_{ik} = \Pr(E_i \mid H_k)$**:** the likelihood of E_i given H_k.
$q_{ik} = \Pr(H_i \mid E_k)$**:** the posterior probability of H_i given evidence E_k.

We have

$$\begin{aligned} \Pr(E_k) &= \sum_i \Pr(E_k \mid H_i) \Pr(H_i) = \sum_i p_i p_{ki}, \\ q_{ik} &= \Pr(H_i \mid E_k) = \frac{\Pr(E_k \mid H_i) \Pr(H_i)}{\Pr(E_k)} = \frac{p_i p_{ki}}{\sum_i p_i p_{ki}}. \end{aligned} \tag{1.1}$$

Bayesian (utility maximazing) rationality mandates that

Without any evidence E_k**:** choose an action A_s such that

$$s \in \operatorname{argmax}_j \sum_i p_i u_{ij} \tag{1.2}$$

With evidence E_k**:** choose an action A_s such that

$$s \in \operatorname{argmax}_j \sum_i q_{ik} u_{ij} \tag{1.3}$$

Date: January 28, 2017.

Seminar on Decision Theory, Lent Term 2016.

With all potential evidence: given the probability of E_k, namely $\Pr(E_k)$, the expected utility of the best action(s) recommended with experiment T is

$$\begin{aligned}\sum_k \Pr(E_k) \max \sum_i q_{ik} u_{ij} &= \sum_k \Big(\sum_i p_i p_{ki} \Big) \max \sum_i \frac{p_i p_{ki}}{\sum_i p_i p_{ki}} u_{ij} \\ &= \sum_k \max \sum_i p_i p_{ki} u_{ij}.\end{aligned} \tag{1.4}$$

Theorem 1.1.

$$\sum_k \max_j \sum_i p_i p_{ki} u_{ij} \geq \max_j \sum_i p_i u_{ij}. \tag{1.5}$$

Proof. Observe that, for any real valued function $f(j,k)$, we have

$$\sum_k \max_j f(j,k) \geq \max_j \sum_k f(j,k).$$

Since $\sum_k p_{ki} = 1$, put $f(j,k) = \sum_i p_i p_{ki} u_{ij}$, we get

$$\sum_k \max_j \sum_i p_i p_{ki} u_{ij} \geq \max_j \sum_k \sum_i p_i p_{ki} u_{ij} = \max_j \sum_i p_i u_{ij}.$$

2. Dilation

Let $\mathcal{P}$ be a set of probability measures that represents the agent's credal states. For any given event E, define the following probabilistic characterisations of E:

$$\begin{aligned}\underline{\Pr}(E) &= \inf \Big\{ \Pr(E) : \Pr \in \mathcal{P} \Big\}, \\ \overline{\Pr}(E) &= \sup \Big\{ \Pr(E) : \Pr \in \mathcal{P} \Big\}, \\ \underline{\Pr}(E \mid F) &= \inf \Big\{ \Pr(E \mid F) : \Pr \in \mathcal{P} \Big\}, \\ \overline{\Pr}(E \mid F) &= \sup \Big\{ \Pr(E \mid F) : \Pr \in \mathcal{P} \Big\}.\end{aligned} \tag{2.1}$$

Definition 2.1. Let $E_1, \ldots, E_n$ be a partition of the state space and let H be an event, say that the partition (strictly) *dilates* H if, for any E_i ($i = 1, \ldots, n$),

$$\underline{\Pr}(H \mid E_i) < \underline{\Pr}(H) \leq \overline{\Pr}(H) < \overline{\Pr}(H \mid E_i). \tag{2.2}$$

Example 2.2. Suppose that a fair coin is tossed twice, in such a way that the second toss may depend on the outcome of the first, but You know nothing about the type or degree of dependence. Let H_1, T_1, H_2, T_2 denote the possible outcomes of the two tosses.

I: Since the coin is known to be fair we have

$$\underline{\Pr}(H_1) = \overline{\Pr}(H_1) = 1/2 \quad \text{and} \quad \underline{\Pr}(H_2) = \overline{\Pr}(H_2) = 1/2.$$

II: To model the ignorance of the degree of dependence we have

$$\underline{\Pr}(H_1 \cap H_2) = 0 \quad \text{and} \quad \overline{\Pr}(H_1 \cap H_2) = \overline{\Pr}(H_1) = 1/2.$$

which are the cases where either the occurrence of H_1 may never lead to H_2 or it may alway do so.

III: Now suppose that You learn that the second toss is head, then you want to encoparate this information in estimating the result of the first toss. By (2.1) we have

$$\begin{aligned}\underline{\Pr}(H_1 \mid H_2) &= \inf\left\{\frac{\Pr(H_1 \cap H_2)}{\Pr(H_2)} : \Pr \in \mathcal{P}\right\} \\ &= 2 \cdot \underline{\Pr}(H_1 \cap H_2) = 0\end{aligned}$$

and

$$\begin{aligned}\overline{\Pr}(H_1 \mid H_2) &= \sup\left\{\frac{\Pr(H_1 \cap H_2)}{\Pr(H_2)} : \Pr \in \mathcal{P}\right\} \\ &= 2 \cdot \overline{\Pr}(H_1 \cap H_2) = 1.\end{aligned}$$

By a symmetric argument, we have $\underline{\Pr}(H_1 \mid T_2) = 0$ and $\overline{\Pr}(H_1 \mid T_2) = 1$. Since H_2 and T_2 partition the outcome space, we then have that

$$0 = \underline{\Pr}(H_1 \mid X) < \underline{\Pr}(H_1) = 1/2 = \overline{\Pr}(H_1) < \overline{\Pr}(H_1 \mid X) = 1. \tag{2.3}$$

where $X = H_2$ or T_2. In other words, Your probabilistic estimation of the first toss *dilates* from $1/2$ to the vacuous unit interval *no matter what you learn about the result of the second toss*.[1] ◁

References

Good, I. J. (1967). On the principle of total evidence. *The British Journal for the Philosophy of Science*, 17(4):319–321.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall New York.

[1] Walley (1991, p.298-299).