Postscript 2005: Uniformity and the Chan-Darwiche Distance

Carl G. Wagner
Department of Mathematics
The University of Tennessee
Knoxville, TN 37996-1300

Abstract

Chan and Darwiche have recently introduced a metric on probabilities with the property that if p is revised to q, and p' to q', based on identical learning, with q' constructed in accord with the Uniformity Rule, then the posteriors q and q' are the same distance from their respective priors.

When Ω is finite the Uniformity Rule has intriguing connections with a distance function on probabilities introduced by Chan and Darwiche (2002). Assume for simplicity that the probabilities p and q are strictly coherent, and defined on all subsets of Ω. The *Chan-Darwiche distance* $CD(p, q)$ is defined by

$$CD(p, q) := \log R - \log r, \tag{1}$$

where

$$R := \max_{\omega \in \Omega} \frac{q(\omega)}{p(\omega)} \quad \text{and} \quad r := \min_{\omega \in \Omega} \frac{q(\omega)}{p(\omega)}. \tag{2}$$

It is straightforward to show that CD is a *metric* on the set of all strictly coherent probability measures on the algebra $\mathbf{A} = 2^{\Omega}$, i.e., that

$$CD(p, q) \geqslant 0, \text{ with } CD(p, q) = 0 \text{ iff } p = q; \tag{3}$$

$$CD(p, q) = CD(q, p); \qquad \text{and} \tag{4}$$

$$CD(p, q) \leqslant CD(p, p') + CD(p', q). \tag{5}$$

The distance $CD(p, q)$ yields uniform bounds on the Bayes factors $\beta_p^q(A : B)$.

Theorem. *For all nonempty events $A, B \in 2^{\Omega}$,*

$$e^{-CD(p,q)} \leqslant \beta_p^q(A:B) \leqslant e^{CD(p,q)}. \tag{6}$$

Proof. Suppose that $\max_\omega q(\omega)/p(\omega)$ and $\min_\omega q(\omega)/p(\omega)$ are attained, respectively, at $\omega = \omega_2$ and $\omega = \omega_1$. Then

$$\frac{q(\omega_1)}{p(\omega_1)} p(\omega) \leqslant q(\omega) \leqslant \frac{q(\omega_2)}{p(\omega_2)} p(\omega). \tag{7}$$

Summing (7) over all $\omega \in A$, and all $\omega \in B$ yields

$$\frac{q(\omega_1)}{p(\omega_1)} \leqslant \frac{q(A)}{p(A)}, \frac{q(B)}{p(B)} \leqslant \frac{q(\omega_2)}{p(\omega_2)}, \tag{8}$$

whence,

$$\frac{q(\omega_1)}{p(\omega_1)} \Big/ \frac{q(\omega_2)}{p(\omega_2)} \leqslant \frac{\pi_p^q(A)}{\pi_p^q(B)} \leqslant \frac{q(\omega_2)}{p(\omega_2)} \Big/ \frac{q(\omega_1)}{p(\omega_1)}, \tag{9}$$

which is equivalent to (6) by (1), (2), and (1.4) of "Commuting Probability Revisions: The Uniformity Rule." □

Note that the bounds in (6) are sharp, the upper bound being attained when $A = \{\omega_2\}$ and $B = \{\omega_1\}$, and the lower bound when $A = \{\omega_1\}$ and $B = \{\omega_2\}$. In particular,

$$\begin{aligned} CD(p,q) &= \max_{A,B \subset \Omega} \log \beta_p^q(A:B) \\ &= \max_{\omega,\omega' \in \Omega} \log \beta_p^q(\{\omega\}:\{\omega'\}). \end{aligned} \tag{10}$$

Hence if p is revised to q, and p' is revised to q', based on identical learning, and we construct q' in accord with the uniformity rule, then $CD(p,q) = CD(p',q')$. *So revisions based on identical learning, carried out according to the dictates of the uniformity rule, move us the same CD-distance from the priors in question.* As can be seen from the elementary example

$$\begin{array}{rcc} & \omega_1 & \omega_2 \\ p: & 1/2 & 1/2 \\ q: & 4/5 & 1/5 \end{array} \qquad \begin{array}{rcc} & \omega_1 & \omega_2 \\ p': & 2/5 & 3/5 \\ q': & 8/11 & 3/11 \end{array} \quad , \tag{11}$$

where $CD(p,q) = CD(p',q') = 2\log 2$, this fails to be the case for other measures of distance, including the *Euclidean distance*

$$ED(p,q) = \left(\sum_{\omega}\big(p(\omega) - q(\omega)\big)^2\right)^{1/2}, \tag{12}$$

the *variation distance*

$$\begin{aligned}\|p - q\| &= \max\{|p(A) - q(A)| : \ A \subset \Omega\} \\ &= \frac{1}{2}\sum_{\omega}|p(\omega) - q(\omega)|,\end{aligned} \tag{13}$$

the *Hellinger distance*

$$H(p,q) = \sum_{\omega}\left(\sqrt{p(\omega)} - \sqrt{q(\omega)}\right)^2, \tag{14}$$

and the *Kullback-Leibler information number*

$$KL(q,p) = \sum_{\omega} q(\omega)\log\big(q(\omega)/p(\omega)\big). \tag{15}$$

In what follows, $CD(p,q)$ is denoted simply by d. Setting $B = \bar{A}$ in (6) yields

$$e^{-d}p(A)\,/p(\bar{A}) \leqslant q(A)\,/q(\bar{A}) \leqslant e^{d}p(A)\,/p(\bar{A}) \tag{16}$$

and since $x \mapsto x/1 + x$ is an increasing function of x, (16) implies that

$$\frac{p(A)e^{-d}}{p(A)\,[e^{-d} - 1] + 1} \leqslant q(A) \leqslant \frac{P(A)e^{d}}{p(A)\,[e^{d} - 1] + 1}. \tag{17}$$

Since $\beta_p^q(A : \bar{A}|B) = \beta_p^q(AB : \bar{A}B)$, it follows from (6), with $A = AB$ and $B = \bar{A}B$, that

$$e^{-d} \leqslant \beta_p^q(A : \bar{A}|B) \leqslant e^{d}, \tag{18}$$

which leads, by an argument similar to the above, to the inequality

$$\frac{p(A|B)e^{-d}}{p(A|B)\,[e^{-d} - 1] + 1} \leqslant q(A|B) \leqslant \frac{p(A|B)e^{d}}{p(A|B)\,[e^{d} - 1] + 1}. \tag{19}$$

Chan and Darwiche [2002, Theorem 2.2] prove (18) and (19) and show that the latter has interesting applications to sensitivity analysis in belief networks. Indeed, it was a desire to identify bounds on conditional odds and conditional probabilities that led these authors to the discovery of the metric CD.

The implications of (16) and (18) for probabilities constructed by the Uniformity Rule are obvious. Suppose that p is revised to q, and p' to q', based on identical learning, with q' constructed in accord with this rule. Then, although there are in general many events A and B for which $\beta_{p'}^{q'}(A : B) \neq \beta_p^q(A : B)$ and $\beta_{p'}^{q'}(A : \bar{A}|B) \neq \beta_p^q(A : \bar{A}|B)$, all four of these quantities have precisely the same lower and upper bounds. The inequalities (17) and (19) have similar, slightly weaker consequences, which we leave to the reader to delineate.

Reference

1. H. Chan and A. Darwiche, A distance measure for bounding probabilistic belief change, Proc. of the 18th Int. Conf. on Artificial Intelligence (AAAI-2002), Menlo Park, CA, AAAI Press, 539–545.