Philosophy of Science (), 1–
doi:

PSA PHILOSOPHY OF SCIENCE ASSOCIATION PHILOSOPHY OF SCIENCE

RESEARCH ARTICLE

Credence and Belief: Distance- and Utility-Based Approaches

Minkyung Wang[1] and Chisu Kim[2]

[1]Munich Center for Mathematical Philosophy, [2]Independent Researcher,
*Corresponding author: Minkyung Wang; minkyungwang@gmail.com

(Received XX XX XXXX; revised XX XX XXXX; accepted XX XX XXXX)

Appendix

A. Remark

In the following remark, we gather some useful results about maximal support, $\triangle_q$, and $\mathbb{F}_p$. These will be used in the proofs in the next section. Recall that $\mathrm{ri}(X)$ is the relative interior of X ($\subseteq \mathbb{R}^m$). Proofs are omitted but can be easily checked.

Remark. Let $q \in \triangle^M$ and $\mathrm{MSupp}(q) = \{w_1, \ldots, w_k\}$.

(i) There is a weighting vector $(\lambda_i)_{i\leq k} \in (0,1]^k$ such that $q = \sum_{i=1}^k \lambda_i v_{w_i}$.
(ii) If $q \in \mathrm{ri}(\triangle^M)$, then $\mathrm{MSupp}(q) = W$ and $\triangle_q = \triangle^M$.
(iii) $\mathrm{ri}(\triangle_q) = \{p \in \triangle^M \mid \mathrm{MSupp}(p) = \mathrm{MSupp}(q)\}$.
(iv) $q \in \mathrm{ri}(\triangle_q)$.
(v) $\mathrm{MSupp}(p) \subseteq \mathrm{MSupp}(q)$ if and only if (iff) $p \in \triangle_q$ iff $q \in \mathbb{F}_p$.
(vi) If $\mathrm{MSupp}(p) \subseteq \mathrm{MSupp}(q)$, then $\mathbb{F}_q \subseteq \mathbb{F}_p$.

B. Proofs

Lemma 1 (Invariance under the same ouput representation (IOR)). *Let B and B' be non-empty subsets of W. If the corresponding points in $\triangle^{\mathscr{F}}$ are the same, i.e., $b = b'$, then, for all $A \in \mathscr{F}$,*

(i) $B \subseteq A$ iff $b_A = 1$ and $B \subseteq A^c$ iff $b_A = 0$, and thus
(ii) $B \subseteq A$ iff $B' \subseteq A$ and $B \subseteq A^c$ iff $B' \subseteq A^c$, where A^c is the complement of A.

Proof. In $\triangle^{\mathscr{F}}$, $w \in A$ iff $(v_w)_A = 1$ for all $w \in W$ and for all $A \in \mathscr{F}$. Thus, $B \subseteq A$ means that for all $w' \in B$, $(v_{w'})_A = 1$, which is equivalent to $b_A \left(= \sum_{w'\in B} 1/|B|(v_{w'})_A\right) = 1$. Similarly, $w \in A^c$ iff $(v_w)_A = 0$ for all $w \in W$ and for all $A \in \mathscr{F}$. Thus, $B \subseteq A^c$ means that for all $w' \in B$, $(v_{w'})_A = 0$, which is equivalent to $b_A = 0$.[1] Since $b = b'$ means that for all $A \in \mathscr{F}$, $b_A = b'_A$, the rest of the claim follows. □

Theorem 1 (Characterization of DM rule). *A binarization rule (BR) G is a distance minimization (DM) rule in $\triangle^M$ iff*

(i) G satisfies invariance under the same input representation (IIR) in $\triangle^M$, and
(ii) G satisfies the suspension principle.

[1]The following is another proof for the first part of Lemma 1. If $B \subseteq A$, then $b_A = U(B)(A) = \sum_{w\in A} U(B)(w) = \sum_{w\in A\cap B} U(B)(w) = \sum_{w\in B} U(B)(w) = 1$. If $B \nsubseteq A$, then $\sum_{w\in A\cap B} U(B)(w) \neq \sum_{w\in B} U(B)(w)$. Thus, $b_A \neq 1$.

Proof. ($\to$) Since G has the form $G(P) = \mathrm{argmin}_b\, d(p,b)$ for some divergence d, (i) and (ii) hold.

($\leftarrow$) Define d as follows: For $b \in U^M$,

$$d(p,b) := \begin{cases} 0 & \text{if } b \in G(P) \text{ and } p = b, \\ 1 & \text{if } b \in G(P) \text{ and } p \neq b, \\ 2 & \text{if } b \notin G(P), \end{cases}$$

and for $q \in \triangle^M \setminus U^M$, $d(p,q) = 0$ if $p = q$, otherwise $d(p,q) = 1$. This is well defined thanks to IIR. And d is a divergence: If $p = b \in U^M$ then $G(P) = b$ by the suspension principle and thus $d(p,b) = 0$. If $p \neq b$, then $d(p,b) = 1$ if $b \in G(P)$, otherwise $d(p,b) = 2$, and thus $d(p,b) > 0$. For $q \in \triangle^M \setminus U^M$, $d(p,q) = 0$ iff $p = q$. Furthermore, we have $G(P) = \mathrm{argmin}_b\, d(p,b)$: If $p = b$, $\mathrm{argmin}_b\, d(p,b) = \{b\} = G(P)$ by the definition of d and the suspension principle. If $p \neq b$, $\mathrm{argmin}_b\, d(p,b) = \{b \mid d(p,b) = 1\} = G(P)$. □

To prove Theorem 2, we need the following lemma, which shows that the directional derivative of a convex function is linear, if it exists and is finite.

Lemma 2 (Linearity of directional derivative)**.** *Let $\Phi\colon \triangle^M \to \mathbb{R}$ be a convex function. The following statements are equivalent:*

(i) For all $p \in \triangle^M$ and $q \in \mathbb{F}_p$, the directional derivative $\nabla_{p-q}\Phi(q)$ exists and is finite.

(ii) For all $q \in \triangle^M$ there exists $f \in \mathbb{R}^m$ such that, for all $p \in \triangle_q$, $\nabla_{p-q}\Phi(q) = f \cdot (p-q)$.

Proof. (ii) $\to$ (i) is straightforward.

(i) $\to$ (ii): Suppose that f is a subgradient of Φ at q. As $q \in \mathrm{ri}(\triangle_q)$, the existence of a subgradient is guaranteed by the convexity of Φ. For $h > 0$, from the definition of subgradient we have that $\Phi(q + h(p-q)) \geq \Phi(q) + f \cdot h(p-q)$ for all $p \in \triangle_q$, that is,

$$\frac{\Phi(q + h(p-q)) - \Phi(q)}{h} \geq f \cdot (p-q).$$

For $h > 0$ small enough that $q - h(p-q) \in \triangle_q$ (such an h exists since $q \in \mathrm{ri}(\triangle_q)$), we have that $\Phi(q - h(p-q)) \geq \Phi(q) - f \cdot h(p-q)$ for all $p \in \triangle_q$, that is,

$$\frac{\Phi(q) - \Phi(q - h(p-q))}{h} \leq f \cdot (p-q).$$

Since Φ has a finite directional derivative at q in the direction of $p - q$, with $h \to 0$ we get $\nabla_{p-q}\Phi(q) = f \cdot (p-q)$.[2] □

Theorem 2 (Representation of DM(Bregman) by strictly proper expected (epistemic) utility maximization (EUM(SP)))**.** *Let D be a Bregman divergence in $\triangle^M$. Then, for all $p, q \in \triangle^M$ and any probability function $P \in \mathbb{P}(W)$ represented by p,*

$$D(p,q) = \mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)],$$

and thus $\mathrm{argmin}_b\, D(p,b) = \mathrm{argmin}_b\, \mathbb{E}_{w\sim P}[D(v_w,b)]$.

Proof. First, assume that $q \in \mathbb{F}_p$. Then, not only the left-hand side but also the right-hand side is finite because

$$\mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)] = \sum_{w \in \mathrm{Supp}(P)} P(w) D(v_w,q) - \sum_{w \in \mathrm{Supp}(P)} P(w) D(v_w,p)$$

and, for all $w \in \mathrm{Supp}(P)$ $(\subseteq \mathrm{MSupp}(p) \subseteq \mathrm{MSupp}(q))$, $D(v_w,q)$ and $D(v_w,p)$ are finite. Let $D(p,q) = \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q)$. Then

$$\begin{aligned} \mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)] &= \mathbb{E}_{w\sim P}[\Phi(v_w) - \Phi(q) - \nabla_{v_w-q}\Phi(q)] \\ &\quad - \mathbb{E}_{w\sim P}[\Phi(v_w) - \Phi(p) - \nabla_{v_w-p}\Phi(p)] \\ &= \Phi(p) - \Phi(q) - \mathbb{E}_{w\sim P}[\nabla_{v_w-q}\Phi(q)] + \mathbb{E}_{w\sim P}[\nabla_{v_w-p}\Phi(p)] \\ &= \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q). \end{aligned}$$

[2] Notice that for any subgradients f and f' of Φ at q, we have $f \cdot (p-q) = f' \cdot (p-q)$ for all $p \in \triangle_q$. This shows the uniqueness of the subgradient of $\Phi \restriction \triangle_q$ at q, which indicates differentiability at q.

The last equality follows from the fact that

$$\mathbb{E}_{w\sim P}[\nabla_{v_w-q}\Phi(q)] = \mathbb{E}_{w\sim P}[f\cdot(v_w-q)] = f\cdot\mathbb{E}_{w\sim P}[(v_w-q)] = f\cdot(\mathbb{E}_{w\sim P}[v_w]-q),$$

where f is a (sub)gradient at q and $\mathbb{E}_{w\sim P}[\vec{g}(w)] = (\mathbb{E}_{w\sim P}[g_i(w)])_{i\le m}$ for $\vec{g}\colon W\to\mathbb{R}^m$. Since $w\in \mathrm{MSupp}(q)$ for all $w\in\mathrm{Supp}(P)$, this holds by the linearity of expectation and the linearity of the directional derivative of a convex function, which is proved by Lemma 2. Our claim holds from $\mathbb{E}_{w\sim P}[v_w]=\sum_{w\in W}P(w)v_w = p$.

Next, assume that $q\notin\mathbb{F}_p$. Then

$$\begin{aligned} D(p,q) = \lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}} D(p,x) &= \lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}} (\mathbb{E}_{w\sim P}[D(v_w,x)] - \mathbb{E}_{w\sim P}[D(v_w,p)]) \\ &= \lim_{\substack{x\to q\\ :x\in\mathbb{F}_p}} \sum_w P(w)D(v_w,x) - \mathbb{E}_{w\sim P}[D(v_w,p)] \\ &= \sum_{w\in\mathrm{Supp}(P)} P(w) \lim_{\substack{x\to q\\ :\mathrm{MSupp}(p)\subseteq\mathrm{MSupp}(x)}} D(v_w,x) - \mathbb{E}_{w\sim P}[D(v_w,p)] \\ &= \mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)]. \end{aligned}$$

The fourth equality holds since $P(w), D(v_w,x)\ge 0$. Let us explain why the last equality holds: For any $w\in\mathrm{Supp}(P)$, thus for any w such that $\{w\}=\mathrm{MSupp}(v_w)\subseteq\mathrm{MSupp}(p)$, $\lim_{x\to q:\,\mathrm{MSupp}(p)\subseteq\mathrm{MSupp}(x)} D(v_w,x)$ exists because $\{x\in\triangle^M \mid \mathrm{MSupp}(p)\subseteq\mathrm{MSupp}(x)\}\subseteq\{x\in\triangle^M\mid\mathrm{MSupp}(v_w)\subseteq\mathrm{MSupp}(x)\}$.[3] Moreover, $\lim_{x\to q:\,\mathrm{MSupp}(p)\subseteq\mathrm{MSupp}(x)} D(v_w,x) = D(v_w,q)$ since $\lim_{x\to q:\,\mathrm{MSupp}(v_w)\subseteq\mathrm{MSupp}(x)} D(v_w,x)=D(v_w,q)$. □

Lemma 3. *Let* $I\colon W\times\triangle^M\to[0,\infty]$ *be a strictly proper score. Then* $I(w,\cdot)$ *is finite in* $\mathbb{F}_{v_w}$.

Proof. Since $\mathbb{E}_{w\sim P}[I(w,p)] < \mathbb{E}_{w\sim P}[I(w,q)]$ for all $q\ne p$, $\mathbb{E}_{w\sim P}[I(w,p)]$ should be finite for all $P\in\mathbb{P}(W)$ and $p\in\triangle^M$ such that p is a representation point P. Thus, for all $w\in\mathrm{Supp}(P)$, $I(w,p)$ is finite for all P represented by p. Therefore, for all $w\in\mathrm{MSupp}(p)$, $I(w,p)$ is finite. □

Lemma 4. *A function* $I\colon W\times\triangle^{\mathscr{F}}\to[0,\infty]$ *satisfies invariant expectation under the same representation (IER) if* I *is a partition-wise score, i.e., there is a partition of* W*, say* $W=A_1\cup\cdots\cup A_k$*, such that (i)* $A_1,\ldots,A_k\in\mathscr{F}$ *and (ii) for all* $i\le k$ *we have, for all* $w,w'\in A_i$ *and* $q\in\triangle^{\mathscr{F}}$*,* $I(w,q)=I(w',q)$.

Proof. Since $\mathbb{E}_{w\sim P}[I(w,q)] = \sum_{w\in W}P(w)I(w,q) = \sum_{i\le k}P(A_i)I(w_i,q)$, where w_i is any world in A_i, we have $\mathbb{E}_{w\sim P}[I(w,q)] = \sum_{i\le m} p_{A_i}I(w_i,q) = \mathbb{E}_{w\sim P'}[I(w,q)]$. □

Lemma 5. *Let* $I\colon W\times\triangle^{\mathscr{F}}\to[0,\infty]$ *be additive, i.e., for all* $w\in W$ *and* $p\in\triangle^{\mathscr{F}}$,

$$I(w,p) = \sum_{A\in\mathscr{F}} I_A((v_w)_A, p_A),$$

where $I_A\colon\{0,1\}\times[0,1]\to[0,\infty]$ *for all* $A\in\mathscr{F}$.

(i) I satisfies IER.

(ii) If I is event-wise strictly proper (E-SP), i.e.,

$$\operatorname*{argmin}_{q_A\in[0,1]} \big(p_A I_A(1,q_A) + (1-p_A)I_A(0,q_A)\big) = \{p_A\}$$

for all $A\in\mathscr{F}$ *and* $p_A\in[0,1]$*, then I is strictly proper.*

Proof. (i) We compute the following:

$$\begin{aligned} \mathbb{E}_{w\sim P}[I(w,q)] = \sum_{w\in W}P(w)I(w,q) &= \sum_{w\in W}P(w)\sum_{A\in\mathscr{F}} I_A((v_w)_A,q_A) \\ &= \sum_{A\in\mathscr{F}}\left(\sum_{w\in A}P(w)I_A(1,q_A) + \sum_{w\notin A}P(w)I_A(0,q_A)\right) \\ &= \sum_{A\in\mathscr{F}}\big(p_A I_A(1,q_A) + (1-p_A)I_A(0,q_A)\big). \end{aligned} \tag{B1}$$

[3]For any $r\in\triangle^M$, in the case where $\mathrm{MSupp}(r)\not\subseteq\mathrm{MSupp}(q)$, $\lim_{x\to q}D(r,x)$ might not exist if we do not impose the condition about the sequence that $\mathrm{MSupp}(r)\subseteq\mathrm{MSupp}(x)$, under which $D(r,x)$ can be defined without using a limit.

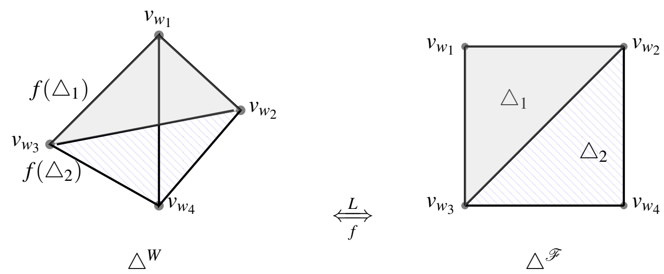

Figure B1: L and f such that $L(f(p)) = p$.

Thus, if $p = p'$, then $\mathbb{E}_{w\sim P}[I(w,q)] = \mathbb{E}_{w\sim P'}[I(w,q)]$.

(ii) The claim follows from B1. □

To prove Theorem 3, we need the following lemmas. Note that we are dealing with continuity not only in $\triangle^W$ but also in $\triangle^{\mathscr{F}}$. The following lemma enables one to find a continuous function assigning a $P \in \mathbb{P}(W)$ to $p \in \triangle^{\mathscr{F}}$.

Lemma A (Continuous selection)**.** *There is a continuous function taking any $p \in \triangle^{\mathscr{F}}$ and giving a $P \in \mathbb{P}(W)$ that is represented by p.*

Proof. Let $|\mathscr{F}| = m$ and $|W| = n$. First, observe that we have a linear function $L\colon \triangle^W \to \triangle^{\mathscr{F}}$ that can be represented by an $m \times n$ binary matrix as follows:

$$\begin{pmatrix} (v_{w_1})_1 & (v_{w_2})_1 & \cdots & (v_{w_n})_1 \\ (v_{w_1})_2 & (v_{w_2})_2 & \cdots & (v_{w_n})_2 \\ \vdots & \vdots & \ddots & \vdots \\ (v_{w_1})_m & (v_{w_2})_m & \cdots & (v_{w_n})_m \end{pmatrix} \begin{pmatrix} P(w_1) \\ P(w_2) \\ \vdots \\ P(w_n) \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix}.$$

Our aim is to find a continuous function $f\colon \triangle^{\mathscr{F}} \to \triangle^W$ satisfying $L(f(p)) = p$ (see figure B1).

First of all, we can triangulate $\triangle^{\mathscr{F}}$ in such a way that $\triangle^{\mathscr{F}}$ is a union of simplexes $\triangle_1, \ldots, \triangle_k$ and $\bigcup_{i=1}^k V(\triangle_i) = V(\triangle^{\mathscr{F}})$, where $V(\triangle)$ denotes the set of all vertexes of a polytope $\triangle$. This is always possible, because $\triangle^{\mathscr{F}}$ is a polytope. For a vertex v of $\triangle^{\mathscr{F}}$ choose one of the omniscient probability measures V_w such that $L(V_w) = v$. This is always possible because for any vertex v we have $\{w \mid v = v_w\} \neq \emptyset$. For any $p \in \triangle_i$ we can uniquely represent p by $p = \sum_{v\in V(\triangle_i)} \lambda_v v$ for some $(\lambda_v)_{v\in V(\triangle_i)}$ such that $\sum_{v\in V(\triangle_i)} \lambda_v = 1$ and $\lambda_v \geq 0$. Then we can define a function f_i from $\triangle_i$ to $\triangle^W$ such that $f_i(p) = \sum_{v\in V(\triangle_i)} \lambda_v L^{-1}(v)$, where $L^{-1}(v)$ denotes the selected omniscient probability measure. Observe that f_i is continuous. Note that for any $q \in \triangle_i \cap \triangle_j$, $f_i(q) = f_j(q)$. Now, we can construct a unique map $f\colon \cup_{i=1}^k \triangle_i \to \triangle^W$ by gluing $f_1, f_2, \ldots, f_k$ where $f \restriction \triangle_i = f_i$ for all $i \leq k$.

Let us check that f is continuous. Suppose that A is a closed subset of $\triangle^W$. Then $f^{-1}(A) = \bigcup_{i=1}^k f_i^{-1}(A)$. Since every $f_i^{-1}(A)$ is closed because of the continuity of f_i, and a finite union of closed sets is closed, it follows that $f^{-1}(A)$ is also closed.

It remains to show that $L(f(p)) = p$ for all $p \in \triangle^{\mathscr{F}}$. First, pick a $\triangle_i$ such that $p \in \triangle_i$. Then

$$L(f(p)) = L\left(f\left(\sum_{v\in V(\triangle_i)} \lambda_v v\right)\right) = L\left(\sum_{v\in V(\triangle_i)} \lambda_v L^{-1}(v)\right) = \sum_{v\in V(\triangle_i)} \lambda_v L(L^{-1}(v)) = \sum_{v\in V(\triangle_i)} \lambda_v v = p,$$

where in the third equality we used the linearity of L. □

From this lemma we see that expected scores are also continuous, as follows.

Lemma B. *Let $I\colon W \times \triangle^M \to [0,\infty]$ be a continuous score, and $p \in \triangle^M$. Let f be a continuous function as in the previous lemma when $\triangle^M = \triangle^{\mathscr{F}}$. When $\triangle^M = \triangle^W$, let f be the identity function. Then $\mathbb{E}_{f(p)}[I(w,p)]$ is continuous at p.*

Proof. $\mathbb{E}_{f(p)}[I(w,p)] = \sum_w f(p)(w)I(w,p)$, and if $w \in \mathrm{Supp}(f(p))$ then $w \in \mathrm{MSupp}(p)$. Thus, for all $w \in \mathrm{Supp}(f(p))$, $I(w,\cdot)$ is finite and continuous at p. Moreover, f is continuous, and the projection on the wth coordinate is continuous. Therefore, our claim holds. □

Theorem 3 (Representation of EUM(SP) by DM(Bregman)). *Let $I \colon W \times \triangle^M \to [0,\infty]$ be a continuous strictly proper score with IER. Then there is a Bregman divergence D in $\triangle^M$ such that, for all $p, q \in \triangle^M$ and any probability function $P \in \mathbb{P}(W)$ represented by p,*

$$D(p,q) = \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)],$$

and thus $\mathrm{argmin}_b\, \mathbb{E}_{w\sim P}[I(w,b)] = \mathrm{argmin}_b\, D(p,b)$.

Proof. For $p, q \in \triangle^M$, let us define a divergence as follows:

$$D(p,q) := \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)].$$

Since I satisfies IER it is well defined, and since I is SP, it is a divergence. We will show that it is a Bregman divergence with $\Phi(p) = -\mathbb{E}_{w\sim P}[I(w,p)]$. Note that Φ is well defined since I satisfies IER.

Claim (i): Φ is continuous, bounded, and strictly convex on $\triangle^M$. By IER and Lemma B, Φ is continuous. Since $I(w,p)$ is finite for all $w \in \mathrm{MSupp}(p)$, it is finite for all $w \in \mathrm{Supp}(P)$ ($\subseteq \mathrm{MSupp}(p)$), and thus $\mathbb{E}_{w\sim P}[I(w,p)]$ is finite. Now let us prove the strict convexity. For $p, q (\in \triangle^M)$ and $\lambda \in (0,1)$ we have

$$\begin{aligned}
-\Phi(\lambda p + (1-\lambda)q) &= \mathbb{E}_{w\sim \lambda P + (1-\lambda)Q}[I(w, \lambda p + (1-\lambda)q)] \\
&= \lambda \mathbb{E}_{w\sim P}[I(w, \lambda p + (1-\lambda)q)] + (1-\lambda)\mathbb{E}_{w\sim Q}[I(w, \lambda p + (1-\lambda)q)] \\
&> \lambda \mathbb{E}_{w\sim P}[I(w,p)] + (1-\lambda)\mathbb{E}_{w\sim Q}[I(w,q)] \\
&= -\lambda\Phi(p) - (1-\lambda)\Phi(q).
\end{aligned}$$

The first equality holds by IER because $\lambda P + (1-\lambda)Q$ is one of the probability distributions that are represented in $\triangle^M$ by $\lambda p + (1-\lambda)q$. The second equality comes from the linearity of expectation, and the inequality in the third line holds because I is SP.

Claim (ii): If $q \in \mathbb{F}_p$, then the directional derivative $\nabla_{p-q}\Phi(q)$ exists and is finite. Moreover $\nabla_{p-\cdot}\Phi(\cdot)$ is continuous at q. Assuming that $q \in \mathbb{F}_p$, we will show that

$$\nabla_{p-q}\Phi(q) = -\mathbb{E}_{w\sim P}[I(w,q)] + \mathbb{E}_{w\sim Q}[I(w,q)]$$

and that it is finite and continuous in q. Note that there is a small enough h such that $q + h(p-q), q - h(p-q) \in \mathrm{ri}(\triangle_q)$ because $p \in \triangle_q$ and $q \in \mathrm{ri}(\triangle_q)$. For $h > 0$,

$$\begin{aligned}
\frac{1}{h}[\Phi(q + h(p-q)) - \Phi(q)] &= -\frac{1}{h}\left[\sum_w (Q + h(P-Q))(w) I(w, q + h(p-q)) - \sum_w Q(w)I(w,q)\right] \\
&= -\frac{1}{h}\sum_w (Q(w) + h(P(w) - Q(w)))[I(w, q + h(p-q)) - I(w,q)] \\
&\quad - \sum_w P(w)I(w,q) + \sum_w Q(w)I(w,q).
\end{aligned}$$

The first equality holds by IER. The last equality holds since every term is finite because

$$\mathrm{Supp}(P), \mathrm{Supp}(Q) \subseteq \mathrm{MSupp}(q + h(p-q)) = \mathrm{MSupp}(q).$$

Since I is strictly proper, we know that

$$\sum_w (Q(w) + h(P(w) - Q(w)))[I(w, q + h(p-q) - I(w,q)] \leq 0,$$

which implies that

$$\frac{1}{h}[\Phi(q + h(p-q)) - \Phi(q)] \geq -\sum_w P(w)I(w,q) + \sum_w Q(w)I(w,q).$$

Similarly, for $h > 0$, we have

$$\frac{1}{h}[\Phi(q) - \Phi(q - h(p-q))] \leq -\sum_w P(w)I(w,q) + \sum_w Q(w)I(w,q).$$

Notice that $\sum_w P(w)I(w,q)$ is continuous in q because for w such that $P(w) \neq 0$, we have $w \in \mathrm{Supp}(P) \subseteq \mathrm{MSupp}(q)$, and thus $I(w,q)$ is continuous in q. By IER and Lemma B, we also have that $\sum_w Q(w)I(w,q)$ is continuous in q. Thismimplies that $\nabla_{p-q}\Phi(q)$ exists as desired. Note that

$$-\sum_w P(w)I(w,q) + \sum_w Q(w)I(w,q) = -\mathbb{E}_{w\sim P}[I(w,q)] + \mathbb{E}_{w\sim Q}[I(w,q)]$$

and it is finite and continuous in q as we indicated above.

Claim (iii): For all $p,q \in \triangle^M$, $D(p,q) = \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q)$ if $q \in \mathbb{F}_p$, otherwise $D(p,q) = \lim_{x\to q\colon x\in\mathbb{F}_p} D(p,x)$, which exists (infinity being allowed as limits). First assume that $q \in \mathbb{F}_p$. By Claim (ii),

$$\begin{aligned}
D(p,q) &= \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)] \\
&= -\mathbb{E}_{w\sim P}[I(w,p)] + \mathbb{E}_{w\sim Q}[I(w,q)] + \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim Q}[I(w,q)] \\
&= \Phi(p) - \Phi(q) - \nabla_{p-q}\Phi(q).
\end{aligned}$$

Otherwise, we need to show that $\lim_{x\to q\colon x\in\mathbb{F}_p} D(p,x) = \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$:

$$\begin{aligned}
\lim_{x\to q\colon x\in\mathbb{F}_p} D(p,x) &= \lim_{x\to q\colon x\in\mathbb{F}_p} \left(\mathbb{E}_{w\sim P}[I(w,x)] - \mathbb{E}_{w\sim P}[I(w,p)]\right) \\
&= \lim_{x\to q\colon x\in\mathbb{F}_p} \sum_w P(w)I(w,x) - \mathbb{E}_{w\sim P}[I(w,p)] \\
&= \sum_{w\in\mathrm{Supp}(P)} P(w) \lim_{x\to q\colon \mathrm{MSupp}(p)\subseteq\mathrm{MSupp}(x)} I(w,x) - \mathbb{E}_{w\sim P}[I(w,p)] \\
&= \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)].
\end{aligned}$$

The third equality holds because $P(w), I(w,x) \geq 0$. The fourth equality holds since, for $w \in \mathrm{Supp}(P)$, $\lim_{x\to q\colon \mathrm{MSupp}(p)\subseteq\mathrm{MSupp}(x)} I(w,x) = I(w,q)$ since $\lim_{x\to q\colon \mathrm{MSupp}(v_w)\subseteq\mathrm{MSupp}(x)} I(w,x) = I(w,q)$. □

Corollary 1. *(i) Let I be a continuous SP score in $\triangle^W$. Then $D(p,q) := \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$ is a Bregman divergence in $\triangle^W$.*

(ii) Let I be a continuous additive E-SP score in $\triangle^{\mathscr{F}}$. Then $D(p,q) := \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$ is an additive Bregman divergence in $\triangle^{\mathscr{F}}$.

Proof. (i) IER always holds in $\triangle^W$.
(ii) Since I is additive, by Lemma 5(i) it has IER and since I is an E-SP score by Lemma 5(ii), it is SP. □

Corollary 2. *(i) Let $I\colon W \times \triangle^M \to [0,\infty]$ satisfy IER. I is continuous SP iff*

$$D_I(p,q) := \mathbb{E}_{w\sim P}[I(w,q)] - \mathbb{E}_{w\sim P}[I(w,p)]$$

is a Bregman divergence.

(ii) Let $D\colon \triangle^M \times \triangle^M \to [0,\infty]$ be a divergence, and suppose that $I_D(w,q) := D(v_w,q)$ satisfies IER. Then D is a Bregman divergence iff

$$D(p,q) = \mathbb{E}_{w\sim P}[D(v_w,q)] - \mathbb{E}_{w\sim P}[D(v_w,p)],$$

and $I_D(w,q)$ is continuous in q.

Proof. (i) ($\rightarrow$) We can easily check this from the proof of Theorem 3.

($\leftarrow$) Since D_I is a divergence, I is SP. Let $p = v_w$ for any $w \in W$. Since $D_I(v_w,q) = I(w,q) - I(w,v_w)$ and D_I is continuous in $\mathbb{F}_{v_w}$, I is continuous in q.

(ii) ($\rightarrow$) We can easily check this from the proof of Theorem 2.

($\leftarrow$) Since D is a divergence, $D(v_w,q)$ is SP. Thus we can apply Theorem 3, and from its proof we know that there is a Bregman divergence d_D that is the same as D. □