TRANSFORMATION OF PROBABILITIES

JULIAN H. BLAU

1. Transformation. Let $\mathcal{R}$ be a $\sigma$-algebra of subsets of $X$, and $\emptyset$ the set of all probability measures $P$ on $\mathcal{R}$. Let $T$ transform $\emptyset$ into itself. For certain sets $E \subseteq \mathcal{R}$, knowledge of $P$ throughout $E$ (i.e., for all subsets of $E$ belonging to $\mathcal{R}$) determines $TP$ throughout $E$. The class of sets having this property will be denoted by $E_T$, or better, since $T$ will be fixed, by $E$. Evidently $E$ contains $0$, $X$, and the complements of atoms. We show that if $E$ is sufficiently large, then $T$ is a linear combination of the identity and a constant. There are applications to the theory of learning and to political theory [1;3;4;6].

Theorem 1. (A) If $E$ contains an algebra $A$ whose Borel extension is $\mathcal{R}$, and if $|\mathcal{R}| > 4$, then $TP = \alpha P + (1 - \alpha)P_0$, where $\alpha \leq 1$ and $P_0 \in \emptyset$. (B) The converse is true with no restriction on $\mathcal{R}$. (C) If $\mathcal{R}$ is infinite, then $\alpha \geq 0$.

In the political interpretation, the elements of $X$ are parties (or political positions). $P$ is the distribution of voters, $T$ is the electoral mechanism, and $TP$ the distribution of seats in the legislature. If $T$ is the identity, the mechanism is Proportional Representation. If $T$ is a constant, the political complexion of the legislature is fixed by law. It will be seen from Theorem 3 that $E \subseteq E$ means that if the complement $-E$ unites in a coalition, the effect is independent of whether this occurs before or after the election. $|\mathcal{R}| > 4$ means essentially that there are more than two parties. Part (A) of the theorem is not true for $|\mathcal{R}| = 4$.

In learning theory, $P$ is a probability distribution of responses, and $TP$ is a new distribution resulting from a learning experience. If $T$ is the identity, there is no learning. If $T$ is a constant, this is one-trial learning.

Bush, Mosteller, and Thompson [4] proved an equivalent theorem for the case $\mathcal{R}$ finite and $E = \mathcal{R}$ (Corollary 3 of Theorem 3). Some of their ideas are used in the proof.

Denote by $B$ the class of sets $E$ such that $P(E) = Q(E)$ implies $TP(E) = TQ(E)$, for all $P, Q \in \emptyset$. The importance of $B$ is that for...
each proper set $E \in B$, there is a function $\gamma_B$ mapping $[0, 1]$ into itself such that $TP(E) = \gamma_B[P(E)]$. We have also $\gamma_B(0) = 0$.

For any class $S$ of subsets of $X$, let $S^*$ denote the class of sets $E$ for which the complement $-E \in S$.

**Proposition 1.** $E \cap E^* \subseteq B$.

**Proof.** Let $E$, $-E \in E$. Let $P(E) = Q(E)$ for two members $P$, $Q$ of $\varnothing$. Define $P'E \in \varnothing$ as follows. For $A \in \mathcal{R}$, let $P'(A) = P(A - E) + A(c)P(E)$, where $c \in E$ and $A(x)$ is the characteristic function of $A$. Define $Q'$ similarly. We have $P' = Q'$ on $E$. Also $P = P'$ and $Q = Q'$ on $-E$. Hence $TP' = TQ'$ on $E$, while $TP = TP'$ and $TQ = TQ'$ on $-E$. In particular, the last two equations are true for $-E$ itself, and, taking complements, also for $E$. We have $TP(E) = TP'(E) = TQ'(E) = TQ(E)$, proving $E \in B$.

Since $A$ is an algebra, $A = A^*$. Thus $B \supseteq E \cap E^* \supseteq A \cap A^* = A$, and so $B \supseteq A$.

Define the set function $u$ on the class $A - \{X\}$ as follows: $u(E) = \gamma_B(0)$. Using the fact that $TP$ is a measure for each $P$, and choosing $P$ so as to vanish on the appropriate sets, it is easy to show that $u$ is a measure on the semiring $A - \{X\}$, and therefore extends uniquely to a measure $u$ on $\mathcal{R}$ [5; 7]. Evidently $u \leq 1$ on $A - \{X\}$, but it would be incorrect to infer that $u(X) \leq 1$.

**Proposition 2.** Let $E \cap F = 0$; $E$, $F$, $E \cup F$ proper sets in $B$; $x$, $y$, $x+y \in [0, 1]$. Then $\gamma_{E \cup F}(x+y) = \gamma_E(x) + \gamma_F(y)$.

**Proof.** Using all the hypotheses, it is easy to show that there is a probability measure $P$ with $P(E) = x$ and $P(F) = y$. For this $P$, 

$$\gamma_{E \cup F}(x+y) = TP(E \cup F) = TP(E) + TP(F) = \gamma_E(x) + \gamma_F(y).$$

If $E$ and $F$ are proper sets in $A$, let $E \sim F$ denote the statement that $\gamma_E(x) - u(E) = \gamma_F(x) - u(F)$ for all $x \in [0, 1]$. $\sim$ is an equivalence relation.

**Proposition 3.** The relation $\sim$ is universal on the proper sets in $A$.

**Proof.** (1) Let $E \subseteq B$, where $\subseteq$ denotes proper inclusion. By Proposition 2, with $F = B - E$ and $y = 0$, $\gamma_B(x) = \gamma_E(x) + \gamma_{B-E}(0)$. Letting $x = 0$, $\gamma_B(0) = \gamma_E(0) + \gamma_{B-E}(0)$. Subtracting, we have $E \sim B$.

(2) If $E \cap F = 0$ and $E \cup F \neq X$, then $E \sim E \cup F \sim F$ by (1).

(3) If $E$, $F$ are incomparable and $E \cap F \neq 0$, then $E \sim E \cap F \sim F$ by (1).

(4) This leaves only the case $F = -E$. For the first time, we invoke the hypothesis $|\mathcal{R}| > 4$, which easily implies $|A| > 4$. Hence $E$ or $F$
must have a proper subset, say $E \supset A$. Then $E \sim A$ by (1) and $A \sim F$ by (2).

In view of Proposition 3 and $|A| > 2$, the equation

$$\gamma(x) = \gamma_E(x) - u(E) \quad \text{for } E \in A - \{0, X\}$$

defines $\gamma(x)$ uniquely. $\gamma$ maps $[0, 1]$ into $[-1, 1]$.

**Proposition 4.** $\gamma(x) \equiv a x$, with $a \leq 1$.

**Proof.** Let $x, y, x + y \in [0, 1]$. Choose $E, F$ so that $E, F, E \cup F$ are proper sets in $A$, and so that $E \cap F = 0$. Here we have used $|\mathbb{R}| > 4$ for the second and last time. By Proposition 2 and the definitions of $u$ and $\gamma$,

$$y(x + y) + u(E \cup F) = y_{E \cup F}(x + y) = y_E(x) + y_{F}(y)$$

$$= y(x) + u(E) + y(y) + u(F).$$

Since $u$ is additive, we conclude that $\gamma(x + y) = \gamma(x) + \gamma(y)$. A bounded function of this type is of the stated form. The proof in [2] can be adapted. Obviously, $a \leq 1$.

Thus $TP(E) = aP(E) + u(E)$ for all $E$ in the semiring $A - \{X\}$. If $a \geq 0$, then $aP + u$ is a measure on $\mathbb{R}$, equal to $TP$ on $A - \{X\}$, and therefore on $\mathbb{R}$. If $a \leq 0$, then $TP - aP$ is a measure on $\mathbb{R}$, equal to $u$ on $A - \{X\}$, and therefore on $\mathbb{R}$. In either case $TP = aP + u$ on $\mathbb{R}$.

In passing, note that

$$1 = TP(X) = a + u(X).$$

If $a = 1$, then $TP \equiv P$. If $a < 1$, define $P_0$ by $(1 - a)P_0 = u$. The main assertion (A) of Theorem 1 follows.

Assertion (B) is immediate, taking $A = \mathbb{R}$. For (C) we require a simple result from set theory. We omit the proof, which is not difficult.

**Lemma.** If $A$ is an infinite algebra of subsets of $X$, then $X$ is the union of a monotone sequence of sets of $A - \{X\}$.

To resume the proof of (C), the infinite cardinality of $\mathbb{R}$ implies the same for $A$. Then we have $u(X) = \lim_{n \to \infty} u(E_n)$ for sets $E_n \subseteq A - \{X\}$. But $u \leq 1$ on $A - \{X\}$, and therefore $u(X) \leq 1$. With (1), we have $a \geq 0$.

For applications to special cases, we need the following closure properties of $E$, which are of independent interest.

**Theorem 2.** (A) If $E, F \subseteq E$, and $E \cup F \neq X$, then $E \cap F \subseteq E$.

(B) $E$ is closed with respect to countable union.
Proof. (A) Let \( P = Q \) on \( E \cap F \). Without loss of generality, assume \( P(E) \leq Q(E) \). Define a new probability measure \( P' \) as equal to \( P \) on \( E \) (i.e., throughout \( E \)), equal to \( Q \) on \( F - E \), and arbitrary on \( X - E - F \) except that \( P'(E) + P'(F - E) + P'(X - E - F) = 1 \). For other sets, \( P' \) is defined by additivity. In verification that the values assigned on \( X - E - F \) are feasible, we observe that this set is not empty and that \( P'(E \cup F) \leq Q(E \cup F) \leq 1 \).

Now \( P = P' \) on \( E \) and \( Q = P' \) on \( F \). Hence \( TP = TP' \) on \( E \) and \( TQ = TP' \) on \( F \). The last two identities are true, therefore, on \( E \cap F \). Hence \( TP = TQ \) on \( E \cap F \), and \( E \cup F \subseteq E \).

We remark that when \( E \cup F = X \), (A) is false in the strong sense that given such overlapping incomparable \( E \), \( F \), there exists a \( T \) for which \( E \) and \( F \) are in \( E_T \), but \( E \cap F \) is not.

(B) Let \( P = Q \) on \( E = \bigcup E_n \), where \( E_n \subseteq E \). Then \( P = Q \) on \( E_n \), which implies \( TP = TQ \) on \( E_n \), for each \( n \). Let \( \{ F_n \} \) be a disjoint sequence having the same partial unions as \( \{ E_n \} \). We have \( TP = TQ \) on \( F_n \), since \( F_n \subseteq E_n \). Then \( TP = TQ \) on \( E \) by countable additivity, and \( E \subseteq E \).

The hypothesis of Theorem 1 may be expressed in two parts:

(I) \( |R| > 4 \), \( E \) contains a class \( S \) whose Borel extension is \( R \), and \( X \subseteq S \).

(II) \( S \) is a ring.

Proposition 5. In Theorem 1, (II) can be weakened to: \( S \) is a semi-ring.

Proof. The class of finite disjoint unions of elements of \( S \) is a ring [5]. Since it contains \( S \), this ring generates \( R \) and contains \( X \). By Theorem 2B, \( E \) contains the ring.

Examples. In all of these, let \( R \) be the class of Borel sets.

(i) \( X \) = the real line. Let \( E \) contain all intervals \([\alpha, \beta) \). (Here and in the following it would suffice to take \( \alpha \) and \( \beta \) rational.) Then \( E \) contains also \([\alpha, \infty) \) and \((-\infty, \alpha) \). With \( 0 \) and \( X \), these finite and semi-infinite intervals constitute a semiring. Proposition 5 applies, and \( TP = \alpha P + (1 - \alpha) P_0 \) with \( 0 \leq \alpha \leq 1 \) as in Theorem 1. This is equally true if \( E \) is assumed instead to contain all proper closed intervals.

(ii) (a) \( X = (0, 1) \), (b) \( X = [0, 1] \), (c) \( X = [0, 1) \). Similar to Example (i).

(iii) \( X \) = Euclidean \( n \)-space \((n > 1) \). Let \( E \) contain all half spaces \( \{ x: x_i \geq \alpha \} \) and \( \{ x: x_i < \alpha \} \). Then \( E \) contains all finite intersections of these sets. (This implication is false for \( n = 1 \).) With \( X \) added, these constitute a semiring, Proposition 5 applies, and \( T \) has the form
stated in Theorem 1. This is true also if \( E \) is assumed to contain all slices \( \{x: x \in [\alpha, \beta]\} \), or alternatively all cells \( \{x: x \in [\alpha_i, \beta_i], i = 1, \ldots, n\} \).

2. Combination. Bush and Mosteller [3] raised the question in learning theory of whether a set \( E \) could be shrunk to a point without making \( T \) ambiguous on the reduced space. More precisely, let \( E \subseteq \mathbb{R} \). A transformation \( C \) of \( \Phi \) into itself is called a combination of \( E \) if it satisfies

\[
(C_1) \quad CP = P \text{ on } -E
\]

and

\[
(C_2) \quad P(E) = Q(E) \quad \text{implies} \quad CP = CQ \text{ on } E.
\]

For example, let \( c \in E \), and let

\[
(2) \quad CP(A) = P(E - A) + A(c)P(E) \quad \text{for each } A \in \mathbb{R}.
\]

We say that \( E \subseteq C \) (\( E \) is combinable) if for each combination \( C \) of \( E \), and for each \( P \in \Phi \), we have

\[
(3) \quad CTP = CTQ.
\]

In learning theory, (3) is called the Combining of Classes condition.

**Theorem 3.** \( C = E^* \).

**Proof.** Let \( E \subseteq C \), and \( C \) be a combination of \( E \). We observe first that (C1) and (C2) imply

\[
(C_3) \quad CP = CQ \text{ if and only if } P = Q \text{ on } -E.
\]

Now let \( P \equiv Q \) on \(-E\). Then \( CP = CQ \). Hence \( CTP = CTCP = CTQ \). Then a second use of (C3) yields \( TP \equiv TQ \) on \(-E\). This proves \( C \subseteq E^* \).

Let \(-E \subseteq E \). Let \( C \) combine \( E \). Then \( P \equiv CP \) on \(-E \), and therefore \( TP \equiv TCP \) on \(-E \). By (C3), \( CTP = CTCP \). Thus \( E^* \subseteq C \).

**Corollary 1.** In Theorem 1, the hypothesis that \( E \) contains the algebra \( A \) can be replaced by \( C \supseteq A \).

**Corollary 2.** (A) The union of two overlapping sets of \( C \) is in \( C \).
(B) \( C \) is closed with respect to countable intersection.

**Corollary 3.** Let \( X \) be finite, \(|X| \geq 2\), let \( R \) be the class of all subsets of \( X \), and \( C = R \). Then \( TP = \alpha P + (1 - \alpha)P_0 \).

This is the Bush-Mosteller-Thompson theorem [4] mentioned.
earlier. Bush and Mosteller [3] showed that $\alpha(|X| - 1) \geq -1$. This bound is attained.

Regarding Example (ii)(c) as the real numbers modulo 1, let $C$ contain all intervals $[\alpha, \beta)$, naturally including the case $\alpha < 1 < \beta$. With 0, this class is a semiring $S$. Since $S = S^*$, also $E \supseteq S$, so that Proposition 5 applies, and $T$ has the familiar form of Theorem 1. In Example (i) (the real line) the corresponding implication is false, even with the additional assumption that $C$ contains all semi-infinite intervals, and similarly for Example (ii). To prove this, we use

**Proposition 6.** Let $T$ be a combination of $E$ of type (2). Then

$$C_T = \left[ \{ \{ c \} \} + \bigcup \{ E \} - \bigcup \{ -E \} \right] \cap R.$$  

(For any subset $S$ of $X$, $\{ S \}$ and $\{ S \}^+$ denote respectively the class of subsets of $S$ and the class of supersets of $S$.) The proof is omitted.

Returning to Example (i), let $T$ be that combination of $E = [c, \infty)$ of type (2) which concentrates $P(E)$ at $c$. We see that $C$ contains all the finite and infinite intervals mentioned above, but that intervals $[\alpha, \beta)$ containing $c$ are not in $E$, and the conclusion of Theorem 1 is false here.

3. **Partition.** A related problem, motivated by learning theory and political theory, is the following. For $n > 1$, let $\varphi_n$ denote the class of all partitions of $X$ into exactly $n$ nonempty parts $X_i$, and let $\mathcal{E}_n$ denote the subclass of partitions (called combinable) for which the $n$-tuple $[P(X_1), \ldots, P(X_n)]$ uniquely determines $[TP(X_1), \ldots, TP(X_n)]$. It is not difficult to show that if each $X_i \in C$, then $(X_1, \ldots, X_n) \in \mathcal{E}_n$. The converse is false, so that the latter statement is actually weaker than the former. Despite this, we have

**Theorem 4.** If $\mathcal{E}_n = \varphi_n$ for some $n < \log_2 |R|$, then $E = C = R$, and $TP = \alpha P + (1 - \alpha) P_0$.

**Proof.** First we show that $E \in E$ for all $E$ divisible into $n - 1$ (proper) parts. We can assume $E \neq X$. Let $P \equiv Q$ on $E$, and let $A \subseteq E$. We can express $A$ as $\bigcup_i A_i$, where either $a = n - 1$ or each $A_i$ is atomic. If $a < n - 1$, then $E - A = \bigcup_{i+1} A_i$ by the hypothesis on $E$. In either case, $P(A_i) = Q(A_i)$ for $i = 1, \ldots, n - 1$ and $P(-E) = Q(-E)$. Hence $TP(A_i) = TQ(A_i)$. Summing from 1 to $a$, $TP(A) = TQ(A)$, and $E \in E$.

Evidently $E \in E$ is proved unless $E$ consists of the union of fewer than $n - 1$ atoms. Let $E$ be the union of $n - 2$ atoms. Since $|R| > 2^n$, $-E$ has three parts, $A, B, C$. Then $E \cup A$ and $E \cup B$ are in $E$, their
union is not $X$, and so their intersection $E$ is in $E$ by Theorem 2A. Similarly for $n - 3$, etc. Thus $E = R$, and the remaining statements follow from Theorems 1 and 3.

The theorem is false for $n \geq \log_2 |R|$.

Let $\sigma_n$ denote the class of partitions of $X$ into $n$ nonempty parts $X_i$, each of which is in a fixed semiring $S$ whose Borel extension is $R$. (The notation is suggested by examples where $S$ consists of intervals.)

When $X$ is the real line, and $S$ the class of intervals $[\alpha, \beta], (-\infty, \alpha], [\alpha, \infty)$, the example at the end of §2 shows that $0 \not\in \sigma_n \subseteq C_n$ for all $n$ does not imply $E = R$. The same is true for $X$ a finite interval. The situation is different for a circle.

**Theorem 5.** Let $X$ be the set of real numbers modulo 1, and $S$ the class of intervals $[\alpha, \beta)$. If $\sigma_n \subseteq C_n$ for some $n$, then $E = C = R$, and $TP = \alpha P + (1 - \alpha) P_0$.

**Proof.** If $n = 2$, then evidently $S \subseteq B$. With the single exception of Proposition 3, the proof of Theorem 1A applies, with the semiring $S$ replacing the algebra $A$. Proposition 1 is superfluous. We show now that the conclusion of Proposition 3 holds also in the present context. All intervals mentioned are proper, i.e., not 0 or $X$.

(1) Let $I_1 \subseteq I$. If $I - I_1$ is an interval, then $I_1 \sim I$ as in Proposition 3. If $I - I_1$ is not an interval, then it is the union of two disjoint intervals $I_2$ and $I_3$. Moreover, $I_1 \cup I_2$ is an interval. Thus $I_1 \sim I_1 \cup I_2 \sim I$.

(2) If $I_1 \cap I_2 = 0$ and $I_1 \cup I_2 \not\subseteq X$, then there is a proper interval $I$ containing $I_1 \cup I_2$. Then $I_1 \sim I \sim I_2$ by (1).

Thus (1) and (2) in the proof of Proposition 3 are true in our present case. (3) and (4) apply unchanged. (This proof that $S \subseteq B$ implies the linearity of $T$ is valid also for $X =$ the real line with $S$ all $[\alpha, \beta)$, and for $X =$ Euclidean $n$-space with $S$ all semiclosed cells.) This completes the proof for $n = 2$.

Next, let $n > 2$. Note that $P \equiv Q$ on $I$ if and only if $P \equiv Q$ for all subintervals of $I$ touching an end point. Hence $I \subseteq E$ if and only if the equality of $P$ and $Q$ for all such subintervals implies the same for $TP$ and $TQ$.

Let $P \equiv Q$ on $I$, and let $I_1$ be a subinterval touching an end point. Write the interval $I - I_1$ as the disjoint union $U_\delta I_1$, and let $I_2 = -I_1$. (Here we have used the fact that $S = S^\ast$.) We have $(I_1, \cdots, I_n) \subseteq C_n$, and $P(I_j) = Q(I_j)$ for all $j$. Hence $TP(I_j) = TQ(I_j)$ for all $j$, and in particular for $j = 1$. Since $I_1$ was arbitrary, this proves $I \subseteq E$. Thus $S \subseteq E$.

Using $S = S^\ast$ again, we have $S \subseteq E \cap E^\ast$. By Proposition 1, $S \subseteq B$, and the first part of the proof applies.
REFERENCES


Antioch College