Some Variations on Variation Independence

A. P. Dawid
Department of Statistical Science
University College London

Abstract

Variation independence of functions is a simple natural 'irrelevance' property arising in a number of applications in Artificial Intelligence and Statistics. We show how it can be alternatively expressed in terms of two other representations of the same underlying structure: equivalence relations and τ-fields.

Key words: Conditional independence; Equivalence relation; Graphoid; Separoid; tau-field; Variation independence

1 AXIOMATIC FRAMEWORK

Let $\mathbf{L}$, with partial order $\leq$, be a lattice. Given a ternary operation $\cdot \perp\!\!\!\perp \cdot \mid \cdot$ on $\mathbf{L}$, we call $(\mathbf{L}, \leq, \perp\!\!\!\perp)$ a *strong separoid* (Dawid 2000), or *abstract graphoid*, if, for all elements x, y, z, w of $\mathbf{L}$:

$$
\begin{array}{lll}
\text{P1}: & x \perp\!\!\!\perp y \mid x & \\
\text{P2}: & x \perp\!\!\!\perp y \mid z & \Rightarrow \quad y \perp\!\!\!\perp x \mid z \\
\text{P3}: & \left.\begin{array}{l} x \perp\!\!\!\perp y \mid z \text{ and} \\ w \leq y \end{array}\right\} & \Rightarrow \quad x \perp\!\!\!\perp w \mid z \\
\text{P4}: & \left.\begin{array}{l} x \perp\!\!\!\perp y \mid z \text{ and} \\ w \leq y \end{array}\right\} & \Rightarrow \quad x \perp\!\!\!\perp y \mid (z \vee w) \\
\text{P5}: & \left.\begin{array}{l} x \perp\!\!\!\perp y \mid z \text{ and} \\ x \perp\!\!\!\perp w \mid (y \vee z) \end{array}\right\} & \Rightarrow \quad x \perp\!\!\!\perp (y \vee w) \mid z \\
\text{P6}: & \left.\begin{array}{l} x \perp\!\!\!\perp y \mid z \text{ and} \\ x \perp\!\!\!\perp z \mid y \end{array}\right\} & \Rightarrow \quad x \perp\!\!\!\perp (y \vee z) \mid (y \wedge z)
\end{array}
$$

It is straightforward to show that, when P1–P5 hold, $x \perp\!\!\!\perp y \mid z$ if and only if $(x \vee z) \perp\!\!\!\perp (y \vee z) \mid z$.

The above properties are abstracted from the notion of conditional independence between random variables (Dawid 1979a; 1980). Indeed, for random variables X, Y, *etc.*, on a probability space $(\Omega, \mathcal{F}, P)$, let $D(X \mid Y)$ denote the distribution, under P, of X given Y, *i.e.* the mapping which associates with each measurable set A in the range of X, its Y-conditional probability $D(X \mid Y)(A) := P(X \in A \mid Y)$ — a function of Y defined up to P-almost sure equality. We say X *is probabilistically independent of* Y *given* Z *(under* P*)*, and write $X \perp\!\!\!\perp_p Y \mid Z\ [P]$, if any of the following properties, easily shown to be equivalent, holds (where equality is to be interpreted as holding P-almost surely, for all sets A, B):

1. $D(X \mid Y, Z) = D(X \mid Z)$.
2. The function $D(X \mid Y, Z)$ of (Y, Z) is equal to a function of Z alone.
3. $D(X, Y \mid Z)(A \times B) = D(X \mid Z)(A) \times D(Y \mid Z)(B)$.

It is then straightforward to verify that the properties P1–P5 hold for $\perp\!\!\!\perp = \perp\!\!\!\perp_p\ [P]$, with $X \leq Y$ interpreted as "X is a function of Y". Property P6 does not hold for probabilistic independence in full generality, but can be obtained under suitable additional conditions (Dawid 1979b).

Once we have abstracted the properties P1–P5 (or P1–P6), they can be seen as supplying a general axiomatic framework for a wide range of concepts of 'irrelevance', not necessarily associated with probability, as well as for other purely mathematical structures (Dawid 1998; 2000).

2 VARIATION INDEPENDENCE

Variation independence is one such irrelevance property, arising in a variety of applications, which is in many ways analogous to — although, indeed, much simpler than — probabilistic conditional independence. Its theory has been most specifically studied in the context of relational databases, where it is closely related to the concept of *embedded multivalued* *dependency* (Sagiv and Walecka 1982; Hill 1993; Wong 1997). It also has important statistical applications in the theory of *meta Markov models* (Dawid and Lauritzen 1993).

Let $\mathbf{F}$ be the set of all functions, with arbitrary range spaces, defined on some non-empty domain space Ω. We do not require any underlying σ-field or probability measure. In the theory of relational databases, one typically takes Ω to be a subset of a Cartesian product space, and confines attention to coordinate projection mappings. However, we do not impose any such restriction here.

Once again, we partially order the set $\mathbf{F}$ by functional contraction: $X \leq Y$ if $Y(\omega) = Y(\eta) \Rightarrow X(\omega) = X(\eta)$. We call two functions X, Y *equivalent*, and write $X \approx Y$, if $X \leq Y$ and $Y \leq X$. Then $(\mathbf{F}, \leq)$ forms a lattice, with join $X \vee Y \approx (X, Y)$

We define the *conditional range of* X*, given* $Y = y$ by: $R(X \mid Y = y) := \{X(\omega) : \omega \in \Omega, Y(\omega) = y\}$. This is non-empty for any $y \in R(Y) := Y(\Omega)$, the *(unconditional) range* of Y. When the conditioning variable Y is clear, we may write simply $R(X \mid y)$. The conditional range represents the residual *logical* uncertainty about X after learning $Y = y$, and is analogous to the conditional distribution $D(X \mid Y = y)$ of X given $Y = y$ in a probabilistic model. This analogy underlies the following concept of 'conditional independence' for this logical framework.

Definition 2.1 We say X *is variation independent of* Y *given* Z *(on* Ω*)*, and write $X \perp\!\!\!\perp_v Y \mid Z\ [\Omega]$ (or, if Ω is understood, just $X \perp\!\!\!\perp_v Y \mid Z$), if any of the following properties, easily shown to be equivalent, holds:

1. For any $(y, z) \in R(Y, Z)$,
$$R(X \mid y, z) = R(X \mid z).$$
2. The function $R(X \mid Y, Z)$ of (Y, Z) is a function of Z alone.
3. For any $z \in R(Z)$,
$$R(X, Y \mid z) = R(X \mid z) \times R(Y \mid z).$$

It may be seen that the property $X \perp\!\!\!\perp_v Y \mid Z\ [\Omega]$ is unaffected if any of the functions in it is replaced by an equivalent function.

It is well-known, and again straightforward to verify, that $(\mathbf{F}, \leq, \perp\!\!\!\perp_v\ [\Omega])$ is a strong separoid.

3 EQUIVALENCE RELATIONS

The set $\mathbf{E}$ of equivalence relations on a set Ω forms a lattice under the partial order of *refinement*: $\sim$ is *less refined than* $\sim'$, written $\sim \leq \sim'$, if $\omega \sim' \eta \Rightarrow \omega \sim \eta$. The corresponding definitions of *join* $\vee$ and *meet* $\wedge$ are given by:

$$\omega\,(\sim \vee \sim')\,\eta \quad \Leftrightarrow \quad \omega \sim \eta \text{ and } \omega \sim' \eta; \tag{1}$$

$$\omega\,(\sim \wedge \sim')\,\eta \quad \Leftrightarrow \quad \left\{ \begin{array}{l} \text{for some } n \geq 1 \text{ there exist} \\ \zeta_1, \ldots, \zeta_n \in \Omega \text{ with} \\ \zeta_1 = \omega,\ \zeta_n = \eta, \text{ and, for} \\ i = 1, \ldots, n-1, \textit{ either} \\ \zeta_i \sim \zeta_{i+1} \textit{ or } \zeta_i \sim' \zeta_{i+1}. \end{array} \right. \tag{2}$$

We define the *composition*, $\sim \circ \sim'$, of $\sim$ and $\sim'$ by:

$$\omega\,(\sim \circ \sim')\,\eta \Leftrightarrow \left\{ \begin{array}{l} \text{there exists } \zeta \in \Omega \text{ with} \\ \omega \sim \zeta \text{ and } \zeta \sim' \eta. \end{array} \right. \tag{3}$$

Typically, this relation is not an equivalence relation. It is readily seen that it will be so if and only if $\sim$ and $\sim'$ *commute*, *i.e.* $\sim \circ \sim' = \sim' \circ \sim$ — a property we denote by $\sim \perp \sim'$. In that case $\sim \circ \sim' = \sim' \circ \sim = \sim \wedge \sim'$.

Definition 3.1 We write $\sim \perp\!\!\!\perp_e \sim' \mid \sim''$ if:

$$(\sim \vee \sim'') \circ (\sim' \vee \sim'') = \sim'' \tag{4}$$

Equivalently, if:

$$(\sim \vee \sim'') \quad \perp \quad (\sim' \vee \sim'') \tag{5}$$
$$\text{and } (\sim \vee \sim'') \wedge (\sim' \vee \sim'') \quad = \quad \sim''. \tag{6}$$

Note that, with this definition, $\sim \perp \sim'$ if and only if $\sim \perp\!\!\!\perp_e \sim' \mid (\sim \wedge \sim')$.

We shall show that $(\mathbf{E}, \leq, \perp\!\!\!\perp_e)$ is a strong separoid, by exhibiting an isomorphism with variation independence.

Let $X \in \mathbf{F}$. We can define a corresponding equivalence relation $\sim_X \in \mathbf{E}$: $\omega \sim_X \omega'$ if and only if $X(\omega) = X(\omega')$. Clearly, two functions are equivalent if and only if they induce the same equivalence relation. Conversely, given any equivalence relation $\sim \in \mathbf{E}$, we can construct the *projection* $\pi_\sim : \Omega \to \Omega/\!\sim$, *i.e.* the function which associates with each $\omega \in \Omega$ the unique $\sim$-equivalence class containing it; and then $\sim$ will be the equivalence relation induced by $\pi_\sim \in \mathbf{F}$.

The above correspondence between equivalence relations and (equivalence classes of) functions may readily be shown to be a lattice isomorphism between $\mathbf{E}$ and $\mathbf{F}$.

Theorem 3.1

$$X \perp\!\!\!\perp_v Y \mid Z \ \Leftrightarrow \sim_X \perp\!\!\!\perp_e \sim_Y \mid \sim_Z.$$

Proof. Note first that $\sim_X \vee \sim_Y = \sim_{(X,Y)}$, *etc.* Thus the property $\sim_X \perp\!\!\!\perp_e \sim_Y \mid \sim_Z$ becomes:

$$\sim_{(X,Z)} \circ \sim_{(Y,Z)} = \sim_Z, \tag{7}$$

or equivalently

$$\sim_{(X,Z)} \perp \sim_{(Y,Z)} \tag{8}$$
$$\text{and } \sim_{(X,Z)} \wedge \sim_{(Y,Z)} = \sim_Z . \tag{9}$$

1. Suppose that $\sim_X \perp\!\!\!\perp_e \sim_Y \mid \sim_Z$. Take $z \in R(Z)$, $x \in R(X|z)$, $y \in R(Y|z)$. Thus there exist $\omega, \eta \in \Omega$ with $X(\omega) = x$, $Y(\eta) = y$, and $Z(\omega) = Z(\eta) = z$. Since $\omega \sim_Z \eta$, by (7) there exists $\zeta \in \Omega$ with $X(\zeta) = X(\omega), = x$; $Y(\zeta) = Y(\eta), = y$; and $Z(\zeta) = z$. This shows that $(x,y) \in R(X,Y|z)$. Thus $X \perp\!\!\!\perp_v Y \mid Z$.

2. Now suppose that $X \perp\!\!\!\perp_v Y \mid Z$. To show (7), it is enough to show that $\omega \sim_Z \eta \Rightarrow \omega\,(\sim_{(X,Z)} \circ \sim_{(Y,Z)})\,\eta$. So suppose $\omega \sim_Z \eta$, so that $Z(\omega) = Z(\eta), = z$, say. Define $x := X(\omega)$, $y := Y(\eta)$. Then $x \in R(X|z)$, $y \in R(Y|z)$ and so, since $X \perp\!\!\!\perp_v Y \mid Z$, $(x,y) \in R(X,Y|z)$. Thus there exists $\zeta \in \Omega$ with $X(\zeta) = x, Y(\zeta) = y, Z(\zeta) = z$. Then we have $\omega \sim_{(X,Z)} \zeta$ and $\zeta \sim_{(Y,Z)} \eta$, whence $\omega\,(\sim_{(X,Z)} \circ \sim_{(Y,Z)})\eta$.

□

Corollary 3.2 *($\mathbf{E}$, $\leq$, $\perp\!\!\!\perp_e$) is a strong separoid.*

4 τ-FIELDS

A class $\mathcal{T}$ of subsets of Ω forms a *τ-field* if it is closed under complementation and arbitrary (not just countable) unions — and thus, also, under arbitrary intersections. The set $\mathbf{T}$ of all τ-fields on Ω is partially ordered by inclusion:

$$\mathcal{T} \leq \mathcal{T}' \Leftrightarrow \mathcal{T} \subseteq \mathcal{T}'.$$

Then $\mathcal{T} \wedge \mathcal{T}' = \mathcal{T} \cap \mathcal{T}'$, while $\mathcal{T} \vee \mathcal{T}'$ is the smallest τ-field containing both $\mathcal{T}$ and $\mathcal{T}'$.

Definition 4.1 We write $\mathcal{T} \perp\!\!\!\perp_t \mathcal{T}' \mid \mathcal{T}''$ if:

$$\left.\begin{array}{l} \alpha \in \mathcal{T} \vee \mathcal{T}'' \\ \beta \in \mathcal{T}' \vee \mathcal{T}'' \\ \alpha \cap \beta = \emptyset \end{array}\right\} \Rightarrow \left\{ \begin{array}{l} \text{there exists } \gamma \in \mathcal{T}'' \\ \text{such that } \alpha \subseteq \gamma, \beta \subseteq \bar{\gamma}. \end{array}\right. \tag{10}$$

When $\mathcal{T}''$ is trivial (in which case we may write $\mathcal{T} \perp\!\!\!\perp_t \mathcal{T}'$) the above condition becomes: $\alpha \in \mathcal{T}$, $\beta \in \mathcal{T}'$, $\alpha \cap \beta = \emptyset \Rightarrow$ either $\alpha = \emptyset$ or $\beta = \emptyset$. This is essentially the relation of 'qualitative independence between σ-fields' due to Rényi (1970). The above extension is based on Bartfai and Rudas (1988).

We shall show that ($\mathbf{T}$, $\leq$, $\perp\!\!\!\perp_t$) is a strong separoid, again indirectly by exhibiting an isomorphism with variation independence.

Let $X \in \mathbf{F}$. Then we can construct an associated τ-field $\mathcal{T}_X \in \mathbf{T}$: $\mathcal{T}_X := \{X^{-1}(B) : B \subseteq R_X\}$. Conversely, given any τ-field $\mathcal{T} \in \mathbf{T}$, we can define a relation $\sim$ by: $\omega \sim \omega'$ if and only if, for all $A \in \mathcal{T}$, $\omega \in A \iff \omega' \in A$. This is easily seen to be an equivalence relation inducing the τ-field $\mathcal{T}$, the equivalence class containing ω being just the intersection of all the sets in $\mathcal{T}$ that contain ω. Let $X \in \mathbf{F}$ be the 'projection function' $\pi_\sim$ associated with $\sim \in \mathbf{E}$. Then $\mathcal{T} = \mathcal{T}_X$.

Again, the above correspondence between τ-fields and (equivalence classes of) functions may be shown to be a lattice isomorphism between $\mathbf{T}$ and $\mathbf{F}$.

Theorem 4.1

$$X \perp\!\!\!\perp_v Y \mid Z \Leftrightarrow \mathcal{T}_X \perp\!\!\!\perp_t \mathcal{T}_Y \mid \mathcal{T}_Z.$$

Proof. We first note that, since $\mathcal{T}_X \vee \mathcal{T}_Y = \mathcal{T}_{(X,Y)}$, the property $\mathcal{T}_X \perp\!\!\!\perp_t \mathcal{T}_Y \mid \mathcal{T}_Z$ becomes:

$$\left.\begin{array}{l} \alpha \in \mathcal{T}_{(X,Z)} \\ \beta \in \mathcal{T}_{(Y,Z)} \\ \alpha \cap \beta = \emptyset \end{array}\right\} \Rightarrow \left\{ \begin{array}{l} \text{there exists } \gamma \in \mathcal{T}_Z \\ \text{such that } \alpha \subseteq \gamma, \beta \subseteq \bar{\gamma}. \end{array}\right. \tag{11}$$

1. Suppose that $\mathcal{T}_X \perp\!\!\!\perp_t \mathcal{T}_Y \mid \mathcal{T}_Z$. Take $z \in R(Z)$, $x \in R(X|z)$, $y \in R(Y|z)$, and let $\alpha := (X,Z)^{-1}(x,z)$, $\beta := (Y,Z)^{-1}(y,z)$. Suppose there existed $\gamma \in \mathcal{T}_Z$ — say $\gamma = Z^{-1}(C)$, and thus $\bar{\gamma} = Z^{-1}(\bar{C})$ — such that $\alpha \subseteq \gamma$, $\beta \subseteq \bar{\gamma}$. Then we would have both $z \in C$ and $z \in \bar{C}$. This contradiction show that no such γ exists whence, since $\mathcal{T}_X \perp\!\!\!\perp_t \mathcal{T}_Y \mid \mathcal{T}_Z$, we must have $\alpha \cap \beta \neq \emptyset$. Taking then $\omega \in \alpha \cap \beta$, we have $X(\omega) = x$, $Y(\omega) = y$, $Z(\omega) = z$, showing that $(x,y) \in R(X,Y|z)$. Thus we have shown $X \perp\!\!\!\perp_v Y \mid Z$.

2. Now suppose that $X \perp\!\!\!\perp_v Y \mid Z$. Take $\alpha \in \mathcal{T}_{(X,Z)}$, $\beta \in \mathcal{T}_{(Y,Z)}$, such that $\alpha \cap \beta = \emptyset$. Let $\gamma := Z^{-1}Z(\alpha)$, $\delta := Z^{-1}Z(\beta)$. Suppose that $\gamma \cap \delta \neq \emptyset$. Take $\xi \in \gamma \cap \delta$, and let $Z(\xi) = z$. Then there exist $\omega \in \alpha, \eta \in \beta$ such that $Z(\omega) = z$, $Z(\eta) = z$. Let $x := X(\omega)$, $y := Y(\eta)$. Then $x \in R(X|z)$, $y \in R(Y|z)$. Since $X \perp\!\!\!\perp_v Y \mid Z$, there exists $\zeta \in \Omega$ with $X(\zeta) = x, Y(\zeta) = y, Z(\zeta) = z$. But then $\zeta \in \alpha \cap \beta$, which is impossible. We deduce that $\gamma \cap \delta = \emptyset$. We thus have $\alpha \subseteq \gamma$, $\beta \subseteq \delta \subseteq \bar{\gamma}$. Since $\gamma \in \mathcal{T}_Z$, we have shown $\mathcal{T}_X \perp\!\!\!\perp_t \mathcal{T}_Y \mid \mathcal{T}_Z$.

$\square$

Corollary 4.2 $(\mathbf{T}, \leq, \perp\!\!\!\perp_t)$ *is a strong separoid.*

It would be pleasant if the above property applied equally if we restricted attention to the set of σ*-fields* in Ω, but this would require further non-trivial conditions.

References

Bartfai, P. and Rudas, T. (1988). Conditionally independent extension of measures. Preprint 57/1988, Mathematical Institute, Hungarian Academy of Sciences.

Dawid, A. P. (1979a). Conditional independence in statistical theory (with Discussion). *Journal of the Royal Statistical Society, Series B*, **41**, 1–31.

Dawid, A. P. (1979b). Some misleading arguments involving conditional independence. *Journal of the Royal Statistical Society, Series B*, **41**, 249–52.

Dawid, A. P. (1980). Conditional independence for statistical operations. *Annals of Statistics*, **8**, 598–617.

Dawid, A. P. (1998). Conditional independence. In *Encyclopedia of Statistical Science (Update Volume 2)*, (ed. S. Kotz, C. B. Read, and D. L. Banks), pp. 146–55. Wiley-Interscience, New York.

Dawid, A. P. (2000). Separoids: A general framework for conditional independence and irrelevance. Research Report 212, Department of Statistical Science, University College London.

Dawid, A. P. and Lauritzen, S. L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. *Annals of Statistics*, **21**, 1272–317.

Hill, J. R. (1993). Comment on Spiegelhalter *et al.* (1993). *Statistical Science*, **8**, 258–61.

Rényi, A. (1970). *Foundations of Probability*. Holden-Day, San Francisco.

Sagiv, Y. and Walecka, S. F. (1982). Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies. *Journal of the Association for Computing Machinery*, **29**, 103–17.

Spiegelhalter, D. J., Dawid, A. P., Lauritzen, S. L., and Cowell, R. G. (1993). Bayesian analysis in expert systems (with Discussion). *Statistical Science*, **8**, 219–83.

Wong, S. K. M. (1997). An extended relational data model for probabilistic reasoning. *Journal of Intelligent Information Systems*, **9**, 181–202.