REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-01-0188

1. REPORT DATE *(DD-MM-YYYY)*
07–2003

2. REPORT TYPE
Technical

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE

OPERATING ON FUNCTIONS WITH VARIABLE DOMAINS

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0601152N

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS

P. G. Calabrese

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SSC San Diego
San Diego, CA 92152-5001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217–5000

10. SPONSOR/MONITOR'S ACRONYM(S)

20090803038

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The sum, difference, product and quotient of two functions with different domains are usually defined only on their common domain. This paper extends these definitions so that the sum and other operations are essentially defined anywhere that at least one of the components is defined. This idea is applied to propositions and events, expressed as indicator functions, to define *conditional* propositions and *conditional* events as three-valued indicator functions that are undefined when their condition is false. Extended operations of "and", "or", "not" and "conditioning" are then defined on these conditional events with variable conditions. The probabilities of the disjunction (or) and of the conjunction (and) of two conditionals are expressed in terms of the conditional probabilities of the component conditionals. Theorem 1 generalizes the standard result that the conditional expectation of the sum of two conditional random variables with disjoint and exhaustive conditions is a weighted sum of the conditional expectations of the component conditional random variables. Because of the extended operations, the theorem is true for arbitrary conditions. Theorem 2 gives a formula for the expectation of the product of two conditional random variables. After the definition of independence of two random variables is extended to accommodate the extended operations, it is applied to the formula of Theorem 2 to simplify the expectation of a product of conditional random variables.

Published in *Journal of Philosophical Logic*, vol. 32, no. 1, February 2003, 1-18.

15. SUBJECT TERMS
Mission Area: Information Science

functions	conditional expectation	three-valued	conditional
random variable	operations	domain	

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UU	18	P. G. Calarese
					19B. TELEPHONE NUMBER *(Include area code)* (619) 553-3680

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

PHILIP G. CALABRESE

OPERATING ON FUNCTIONS WITH VARIABLE DOMAINS *

Received 15 February 2002; received in revised form 10 July 2002

ABSTRACT. The sum, difference, product and quotient of two functions with different domains are usually defined only on their common domain. This paper extends these definitions so that the sum and other operations are essentially defined anywhere that at least one of the components is defined. This idea is applied to propositions and events, expressed as indicator functions, to define *conditional* propositions and *conditional* events as three-valued indicator functions that are undefined when their condition is false. Extended operations of "and", "or", "not" and "conditioning" are then defined on these conditional events with variable conditions. The probabilities of the disjunction (or) and of the conjunction (and) of two conditionals are expressed in terms of the conditional probabilities of the component conditionals. In a special case, these are shown to be weighted averages of the component conditional probabilities where the weights are the relative probabilities of the various conditions. Next, conditional random variables are defined to be random variables X whose domain has been restricted by a condition on a second random variable Y. The extended sum, difference, product and conditioning operations on functions are then applied to these conditional random variables. The expectation of a random variable and the conditional expectation of a conditional random variable are recounted. Theorem 1 generalizes the standard result that the conditional expectation of the sum of two conditional random variables with disjoint and exhaustive conditions is a weighted sum of the conditional expectations of the component conditional random variables. Because of the extended operations, the theorem is true for arbitrary conditions. Theorem 2 gives a formula for the expectation of the product of two conditional random variables. After the definition of independence of two random variables is extended to accommodate the extended operations, it is applied to the formula of Theorem 2 to simplify the expectation of a product of conditional random variables. Two examples end the paper. The first concerns a work force of n workers of different output levels and work shifts. The second example involves two radars with overlapping surveillance regions and different detection error rates. One radar's error rate is assumed to be sensitive to fog and the other radar's error rate is assumed to be sensitive to air traffic density. The combined error rate over the combined surveillance region given heavy fog and moderate air traffic is computed.

KEY WORDS: conditional, conditional expectation, domain, functions, operations, random variable, three-valued

* This research was supported by the U.S. Office of Naval Research via the In-house Laboratory Independent Research program (ILIR).

1. Introduction

From elementary mathematics, we are all familiar with the definitions of the operations of addition, subtraction, multiplication, and division for real-valued functions defined on a common domain D. For each domain element x, the sum function, $(f+g)$, is simply assigned the value $(f+g)(x) = f(x) + g(x)$, the sum of the values of f and g at x, and similarly for the other operations. However, function division, (f/g), requires an extra condition, namely that $g(x)$ not be zero, so that the division can be performed. So (f/g) is said to be "undefined" for any domain values x for which $g(x) = 0$. Thus already the division operation on functions generates new functions that have *restricted* domains, and in general such divisions will generate functions having *different* domains of definition. This leads to the standard definition of operations on functions whose domains are different: If f and g are defined on D and E respectively, then the sum function $(f+g)$ is defined on the intersection of D and E as follows:

$$(1) \qquad (f+g)(x) = \begin{cases} f(x)+g(x) & \text{if } x \in D \cap E, \\ \text{Undefined} & \text{if } x \notin D \cap E. \end{cases}$$

The difference, product and division functions $f-g$, $f*g$, and f/g are similarly defined when D and E are the domains of f and g respectively, but again the quotient (f/g) is also undefined on any zeros of g. Since a summing of f and g cannot be performed for a given domain value x unless both f and g are defined at x, this has seemed to be a reasonable definition, and there has been no reason offered to do it in any other way.

2. Extended Operations of Sum, Difference and Product on Real-Valued Functions

However, recent developments in conditional event algebra [2, 5] suggest that there is good reason for expanding the domain of the sum function to include all values of x that are in at least one of the two domains. Using the set theory notation D' to denote the complement of D, the definition of the sum function $(f+g)$ can advantageously be extended to:

$$(2) \qquad (f+g)(x) = \begin{cases} f(x)+g(x) & \text{if } x \in D \cap E, \\ f(x) & \text{if } x \in D \cap E', \\ g(x) & \text{if } x \in D' \cap E, \\ \text{Undefined} & \text{if } x \in D' \cap E'. \end{cases}$$

In other words, here the sum function takes the value of $f(x)$ if $g(x)$ is undefined, and takes the value $g(x)$ if $f(x)$ is undefined. It then agrees with the old definition on the restricted domain $D \cap E$ and is undefined only on the region outside of $D \cup E$. The other operations on functions, $(f-g)$, $(f*g)$ and (f/g) can be similarly defined. The product $(f*g)$ is completely analogous to the sum with $*$ in place of $+$. The difference is:

$$(3) \qquad (f-g)(x) = \begin{cases} f(x)-g(x) & \text{if } x \in D \cap E, \\ f(x) & \text{if } x \in D \cap E', \\ -g(x) & \text{if } x \in D' \cap E, \\ \text{Undefined} & \text{if } x \in D' \cap E'. \end{cases}$$

The quotient is analogous to the difference:

$$(4) \qquad (f/g)(x) = \begin{cases} f(x)/g(x) & \text{if } x \in D \cap E \text{ and } g(x) \neq 0, \\ f(x) & \text{if } x \in D \cap E', \\ 1/g(x) & \text{if } x \in D' \cap E \text{ and } g(x) \neq 0, \\ \text{Undefined} & \text{if } x \in D' \cap E' \text{ or } g(x) = 0. \end{cases}$$

Note that although it is possible in the sum case, for example, to redefine the two functions f and g to be zero instead of undefined and thereby eliminate the need for the extended operations, a subsequent desire to take the product instead of the sum would require another redefinition. Other advantages are exhibited below.

3. Propositions, Events and Indicator Functions

These kinds of extended definitions have been shown [2, 4, 5] to be useful when defining Boolean-like operations on uncertain conditional propositions or conditional events whose conditions are different.

In a similar vein restricted indicator functions, and their closure under finite addition, have been used successfully by Suppes and Zanotti [13, p. 10], or [12] to define a qualitative relation between pairs of events characterizing the conditional probabilities and conditional expectations of such pairs of events and allowing comparison of conditional probabilities or conditional expectations even when the conditions on the events are different.

For Boolean propositions or events A and B we have familiar and standard operations of "and" ($\wedge$), "or" ($\vee$), and "not" ($\neg$) corresponding to multiplication ($*$), summation ($+$), and negation ($-$) respectively, and also corresponding to intersection ($\cap$), union ($\cup$), and complement ($'$) in the event interpretation. There has been no standard definition of division of propositions or events but now "conditioning" has been recognized to be division. See, for instance, [6].

A proposition A can be represented as a measurable indicator function f_A defined on the universe Ω and taking the value 1 for $\omega \in A$ and 0 for $\omega \in A'$:

$$(5) \qquad f_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \in A'. \end{cases}$$

With this representation, the standard operations on propositions or events can be expressed in terms of function operations. For instance the negation ($'$) of an event A, which is simply the function that is 0 on A and 1 on A', can be expressed as $(f_\Omega - f_A)$ the universal proposition minus f_A. The disjunction ($\vee$) of two propositions f_A and f_B defined respectively on domains A and B is the indicator function $f_{A\vee B}$ defined by

$$(6) \qquad f_{A\vee B}(\omega) = \begin{cases} 1, & \omega \in (A \cup B), \\ 0, & \omega \in (A \cup B)'. \end{cases}$$

This disjunction ($\vee$) of two propositions A and B can be expressed as the maximum $\max(f_A, f_B)$ of the two indicator function f_A, and f_B. Similarly conjunction ($\wedge$) is $\min(f_A, f_B)$. For notational simplicity, a proposition f_A will be denoted simply as "A" but will retain the indicator function meaning.

The probability $P(f_A)$ of a proposition or event f_A is defined to be $P(A)$, the probability of the P-measurable event A on which f takes the value 1. So $P(f_A) = P(f_A = 1) = P(f_A^{-1}(1)) = P(\{\omega \in \Omega : f_A(\omega) = 1\})$.

4. Conditional Propositions, Events and Restricted Indicator Functions

Following De Finetti [8] a conditional $(A|B)$, "A given B" or "A if B", is an ordered pair of propositions or events with three possible truth states: $(A|B)$ takes the truth value of A when B is true but $(A|B)$ is "undefined" or "inapplicable" when B is false. That is,

$$(7) \qquad (A|B) \text{ is } \begin{cases} \text{true} & \text{if } A \text{ and } B \text{ are true,} \\ \text{false} & \text{if } A \text{ is false and } B \text{ is true,} \\ \text{Undefined} & \text{if } B \text{ is false.} \end{cases}$$

While De Finetti's 3-valuedness for conditionals is followed, the interpretation here of the third truth value as "undefined" or "inapplicable" differs markedly from that of De Finetti, who interprets the third truth value as "unknown" and so therefore as something similar or equivalent to a probability value between 0 (false) and 1 (true). By contrast the "inapplicable" interpretation is not a truth or falsity value; it is an indicator of irrelevance. This crucial difference in interpretation leads to a difference in operations.

A conditional can be represented as a restricted (partially defined) indicator function, $(A|B)$:

$$(A|B)(\omega) = \begin{cases} 1 & \text{if } \omega \in A \cap B, \\ 0 & \text{if } \omega \in A' \cap B, \\ \text{Undefined} & \text{if } \omega \in B' \end{cases}$$

$$(8) \qquad = \begin{cases} A(\omega) & \text{if } \omega \in B, \\ \text{Undefined} & \text{if } \omega \notin B. \end{cases}$$

Since $B(\omega) = 1$ if $\omega \in B$, the latter can be expressed as

$$(9) \qquad (A|B)(\omega) = \begin{cases} A(\omega) \wedge B(\omega) & \text{if } \omega \in B, \\ \text{Undefined} & \text{if } \omega \notin B. \end{cases}$$

So $(A|B)$ is just the indicator function $(A \wedge B)$ restricted to the instances $\omega \in B$.

For any conditional $(A|B)$ with $P(B) \neq 0$, the conditional probability $P(A|B)$ is defined as usual to be $P(A \wedge B)/P(B)$. With this definition, the conditionals $(A|B)$ have conditional probabilities that also satisfy the 6 qualitative axioms of Suppes and Zanotti [12] or [13] for a conditional probability measure.

5. Extended Operations on Conditional Propositions

We can expand the definition of a conditional to include cases in which A and B themselves are conditionals. To do this we need only decide on the definition of a conditional whose premise is undefined (U), the other cases being already determined. We will interpret an undefined condition to mean that there is no additional restriction imposed by it:

$$\begin{aligned} [(A|B) \mid (C|D)](\omega) &= [(A|B)(\omega) \mid (C|D)(\omega)] \\ &= \begin{cases} (A|B)(\omega) & \text{if } (C|D)(\omega) \neq 0, \\ \text{Undefined} & \text{if } (C|D)(\omega) = 0 \end{cases} \\ &= \begin{cases} (A|B)(\omega) & \text{if } \omega \in C \vee D', \\ \text{Undefined} & \text{if } \omega \notin C \vee D' \end{cases} \\ (10) \qquad &= \begin{cases} A & \text{if } \omega \in B \wedge (C \vee D'), \\ \text{Undefined} & \text{if } \omega \notin B \wedge (C \vee D'). \end{cases} \end{aligned}$$

So

$$(11) \qquad (A|B) \mid (C|D) = (A \mid B(C \vee D')).$$

With the definition of a conditional event, and using the extended definitions of the operations on functions, definitions can be developed for disjunction ($\vee$), conjunction ($\wedge$) and negation ($'$) to go along with division ($|$) as follows. (Also see [4]).

$$
\tag{12}
\begin{aligned}
[(A|B) \vee (C|D)](\omega) &= (A|B)(\omega) \vee (C|D)(\omega) \\
&= \begin{cases} (A(\omega) \wedge B(\omega)) \vee (C(\omega) \wedge D(\omega)) & \text{if } x \in B \cup D, \\ \text{Undefined} & \text{if } x \notin B \cup D. \end{cases}
\end{aligned}
$$

The latter expression is just the conditional $((A \wedge B) \vee (C \wedge D) \mid (B \vee D))$. So

$$
\tag{13}
(A|B) \vee (C|D) = ((AB \vee CD) \mid (B \vee D)).
$$

Here, juxtaposition of events A and B has replaced the conjunction notation $A \wedge B$. $(A|B) \vee (C|D)$ is just $(AB \vee CD)$ restricted to $(B \vee D)$.

For example, consider the experiment of rolling an ordinary 6-sided die once, and observing the number n showing up on the die. Suppose a wager is made that "if n is even then it will be a 2, or if $n < 5$ then $n < 4$". Each of the two component conditionals is applicable on a different subset of outcomes of the die roll, and combining them with "or" results in a disjunction of two conditional propositions.

By using (13) this disjunction is equivalent to a single conditional, with a conditional probability: $(n = 2 \mid n \text{ is even}) \vee (n < 4 \mid n < 5) = ((n = 2) \vee (n < 4) \mid (n \neq 5)) = (\{1, 2, 3\} \mid \{1, 2, 3, 4, 6\})$, which is the conditional event that if the roll is not 5 then it will be 1, 2 or 3. This has conditional probability $3/5$. By brute force examination of the 6 outcomes, this result can be seen to be consistent with intuition: Only a non-5 is applicable to at least one of the two component conditionals. So a "5" roll doesn't count. Given a non-5 roll the set $\{1, 2, 3\}$ corresponds to winning the wager since "1" and "3" satisfy the second component while "2" satisfies the first component, but "4" and "6" satisfy neither component.

Similarly, for conjunction ($\wedge$)

$$
\begin{aligned}
&[(A|B) \wedge (C|D)](\omega) \\
&\quad = (A|B)(\omega) \wedge (C|D)(\omega) \\
&\quad = \begin{cases} (ABCD)(\omega) & \text{if } \omega \in B \cap D, \\ (AB)(\omega) & \text{if } \omega \in B \cap D', \\ (CD)(\omega) & \text{if } \omega \in B' \cap D, \\ \text{Undefined} & \text{if } \omega \in B' \cap D' \end{cases} \\
&\quad = \begin{cases} (ABCD)(\omega) & \text{if } \omega \in B \cap D, \\ (ABD')(\omega) & \text{if } \omega \in B \cap D', \\ (B'CD)(\omega) & \text{if } \omega \in B' \cap D, \\ \text{Undefined} & \text{if } \omega \in B' \cap D' \end{cases}
\end{aligned}
$$

$$(14) \qquad = \begin{cases} (ABCD \vee ABD' \wedge B'CD)(\omega) & \text{if } x \in B \cup D, \\ \text{Undefined} & \text{if } x \notin B \cup D. \end{cases}$$

So

$$(15) \qquad (A|B) \wedge (C|D) = (ABCD \vee ABD' \vee B'CD) \mid (B \vee D).$$

The negation operation is $[(A|B)'](\omega) = [(A|B)(\omega)]' = A'(\omega)$ if $\omega \in B$, or undefined if $\omega \in B'$. So

$$(16) \qquad (A|B)' = (A'|B).$$

This algebra of uncertain conditional events or propositions has been extensively developed in [2–7] including a theory of deduction for uncertain conditionals extending Boolean deduction. See [2, p. 227] for an account of the Boolean properties retained and lost in the algebra of conditionals.

Concerning the structure of this algebra of conditionals, the conjunction and disjunction operations are obviously commutative and idempotent. They are less obviously also associative. The inapplicable conditional $(1|0)$ is the unique absolute unit since for all conditionals $(x|y)$, $(x|y) \wedge (1|0) = (x|y)$ and $(x|y) \vee (1|0) = (x|y)$. While there is a unique *relative* complement $(a'|b)$ for each conditional $(a|b)$ such that $(a|b) \vee (a'|b) = (0|b)$ and $(a|b) \wedge (a'|b) = (1|b)$, there are no absolute complements. Although $(x|y) \wedge 0 = 0$ and $(x|y) \vee 1 = 1$, it is also true that $(x|y) \wedge 1 = xy$ and $(x|y) \vee 0 = xy$. Neither distributive law holds in general. However, conjunction distributes over disjunction if and only if whenever the outside conditional is true and one of the inside conditionals is false, then the other inside conditional is applicable. Similarly, disjunction distributes over conjunction if and only if whenever the outside conditional is false and one of the inside conditionals is true then the other inside conditional is applicable. (A proof of these facts about distributivity will be provided in a subsequent paper.)

6. Weighted Averages

Among the interesting properties of these operations are the following weighted average formulas ([5, p. 1682]) for the probabilities of the compound conditionals of Equations (13) and (15):

$$\begin{aligned} &P((A|B) \vee (C|D)) \\ &\quad = P(B \mid B \vee D)P(A|B) \\ (17) \qquad &\qquad + P(D \mid B \vee D)P(C|D) - P(ABCD \mid B \vee D), \end{aligned}$$

$$
\begin{aligned}
&P((A|B) \wedge (C|D)) \\
&\quad = P(B \mid B \vee D)P(AD'|B) \\
(18) \qquad &\quad + P(D \mid B \vee D)P(CB'|D) + P(ABCD \mid B \vee D).
\end{aligned}
$$

The last term of (17) and (18) can be written as $P(BD \mid B \vee D)P(AC|BD)$.

If the truth of each conditional implies the inapplicability of the other conditional, that is if AB is a subset of D' and CD is a subset of B', as for example when the two conditions, B and D, are disjoint, then both (17) and (18) reduce to:

$$
\begin{aligned}
&P((A|B) \vee (C|D)) \\
&\quad = P((A|B) \wedge (C|D)) \\
(19) \qquad &\quad = P(B \mid B \vee D)P(A|B) + P(D \mid B \vee D)P(C|D).
\end{aligned}
$$

In any case, without any extra assumptions the following logical equation always holds:

$$(20) \qquad (A|B) \vee (C|D) = (B \mid B \vee D)(A|B) \vee (D \mid B \vee D)(C|D).$$

The right-hand side of (19) is a weighted average of the conditional probabilities of $(A|B)$ and $(C|D)$ where the weights are the probabilities of B and of D given either occurs. Because the conditional expectation of a restricted indicator function of an event A equals the conditional probability of A given the restriction B, this formula is equivalent to the one displayed by Suppes and Zanotti [12, p. 165] or [13, p. 13] for the expectation of the disjunction of restricted indicator functions.

Note that because it is a weighted average the right-hand side of (19) will in general lie between $P(A|B)$ and $P(C|D)$ not above both. So disjunction of conditional events is not always monotonic; $P((A|B) \vee (C|D))$ can be less than $P(A|B)$. Similarly $P((A|B) \wedge (C|D))$ can be greater than $P(A|B)$. This is not strange because in general disjunction or conjunction of a conditional $(A|B)$ with another conditional $(C|D)$ expands the context to $(B \vee D)$, which allows for greater or lesser probability than before the application of the operation: If $(C|D)$ is $(0|\Omega)$ then disjunction with $(A|B)$ yields AB, whose generally lower probability than $P(A|B)$ is $P(A|B)P(B)$. If $(C|D)$ is $(1|\Omega)$ then disjunction with $(A|B)$ yields $(1|\Omega)$, with probability 1.

As a simple example of this non-monotonicity, let $\Omega = \{1, 2, 3, 4, 5, 6\}$, the numbered faces of a 6-sided die thrown once, and let $B = \{2, 4, 6\}$ and $A = \{2, 4\}$. The conditional probability of rolling 2 or 4, given the roll is an even number, equals $2/3$. That is, $P(A|B) = 2/3$. Now suppose also that $C = \{1\}$ and $D = \{1, 3, 5\}$. So $P(C|D) = 1/3$. That is the probability of rolling a 1, given the roll is an odd number, equals $1/3$. Now what is the probability of "rolling a 2 or 4 given the roll is even, or rolling 1 given the roll is odd"? That is, $P((A|B) \vee (C|D)) = ?$ The answer is $P(AB \vee CD \mid B \vee D) = P\{1, 2, 4\}/P(\text{even or odd}) = 3/6 = 1/2$. So here $P((A|B) \vee (C|D))$ is less than $P(A|B)$ alone.

7. Extended Operations on Random Variables

Having extended the operations for functions with different domains and having applied them to extend the operations for conditional propositions, it is possible to extend the operations on random variables and conditional random variables. A real-valued random variable X is a function from a sample space Ω of a probability space $(\Omega, \mathcal{B}, P)$ into the real numbers such that for any real number x, the set of instances $\omega \in \Omega$ for which $X(\omega) < x$ is a member of $\mathcal{B}$, and so has a probability $P\{\omega \in \Omega : X(\omega) < x\}$. It follows that there is a probability that X takes a value in any of the collection of Borel subsets of real numbers, consisting of those subsets that are a countable collection of intersections or unions of the intervals $(-\infty, x)$ or their complements $[x, \infty)$, for any real number x. Of course any interval (x, y) of real numbers is a Borel set.

As with functions, just doing division on random variables in general produces new ones with different domains whenever the divisor assumes the real value 0. Subsequently, using standard techniques, operating with this restricted variable will propagate its restricted domain. However using these extended operations, the domains of functions can be expanded as well as restricted.

While the ordered pair $(A|B)$ for events A, B, is defined and interpreted as "event A given event B is true", the corresponding construction $(X|Y)$, where X and Y are random variables, can not be immediately interpreted because "given Y is true" does not make sense for real-valued random variables. The condition must be an event such as $Y \in B$, the event that Y takes a value in a Borel set of real numbers B.

8. Conditional Random Variables

Let X, Y, W, Z be real-valued random variables on a probability space $\mathcal{P} = (\Omega, \mathcal{B}, P)$ and let A, B, C, $D, \ldots$ be Borel sets on the real line. A conditional random variable $(X \mid Y \in B)$ is just the random variable X restricted to the instances ω for which $Y(\omega) \in B$. That is,

$$(21) \qquad (X \mid Y \in B) = \begin{cases} X(\omega) & \text{if } Y(\omega) \in B, \\ \text{Undefined} & \text{if } Y(\omega) \notin B \end{cases} = X \text{ on } Y^{-1}(B).$$

If $Y^{-1}(B)$ is empty, then $(X \mid Y \in B)$ is completely undefined, defined for no instances ω.

Although conditional probability distributions, conditional density functions and conditional expectations have standard definitions (See, for instance, [11]), the operations of summation, difference, multiplication and division on conditional random variables are all expressed in terms of probability distributions rather than directly. But now these can follow directly from the extended definitions for operations on functions.

9. Operations on Conditional Random Variables

Using the extended definitions for operations on real-valued functions, extended operations for random variables can be defined as follows:

$$
\begin{aligned}
&[(X \mid Y \in B) + (W \mid Z \in D)](\omega) \\
(22) \qquad &= \begin{cases} X(\omega) + W(\omega) & \text{if } Y(\omega) \in B \text{ or } Z(\omega) \in D, \\ \text{Undefined} & \text{if } Y(\omega) \notin B \text{ and } Z(\omega) \notin D. \end{cases}
\end{aligned}
$$

Replacing "+" in (18) with negation $(-)$, or multiplication $(*)$ yields the corresponding operations on the two conditional random variables. Division requires a separate formula due to possible division by zero:

$$
\begin{aligned}
&[(X \mid Y \in B) \div (W \mid Z \in D)](\omega) \\
&= \begin{cases} X(\omega) \div W(\omega) & \text{if } Y(\omega) \in B \text{ or } Z(\omega) \in D, \\ \text{Undefined} & \text{if } Y(\omega) \notin B \text{ and } Z(\omega) \notin D \end{cases} \\
(23) \qquad &= \begin{cases} X(\omega)/W(\omega) & \text{if } Y(\omega) \in B \text{ and } 0 \neq Z(\omega) \in D, \\ X(\omega) & \text{if } Y(\omega) \in B \text{ and } Z(\omega) \notin D, \\ 1/W(\omega) & \text{if } Y(\omega) \notin B \text{ and } 0 \neq Z(\omega) \in D, \\ \text{Undefined} & \text{if } Y(\omega) \notin B \text{ and } (Z(\omega) = 0 \\ & \text{or } Z(\omega) \notin D). \end{cases}
\end{aligned}
$$

10. Expectations and Conditional Expectations

The expectation or average $E(X)$ of a random variable X is defined to be just the sum of the values of X each weighted by its probability. Keeping to an elementary formulation for simplicity of exposition, assume the universe Ω is finite or countable. Then

$$
(24) \qquad E(X) = \sum_{\omega \in \Omega} X(\omega) P(\omega).
$$

If X and W are two random variables defined on Ω then easily $E(X + W) = E(X) + E(W)$.

By standard definitions, the conditional expectation $E(X \mid Y \in B)$, where Y is a random variable on Ω and B is a Borel subset of real numbers, is defined to be

$$\begin{aligned} E(X \mid Y \in B) &= E(X \text{ on } Y^{-1}(B)) \\ &= \sum_{\omega \in \Omega} X(\omega) P(\omega \mid Y^{-1}(B)) \\ &= \sum_{\omega \in \Omega} X(\omega) P(\omega \wedge Y^{-1}(B)) / P(Y^{-1}(B)) \\ &= [1/P(Y \in B)] \sum_{Y(\omega) \in B} X(\omega) P(\omega). \qquad (25) \end{aligned}$$

Note here that if $P(Y^{-1}(B)) = 0$, then the conditional expectation is undefined. Otherwise, $P(\omega \mid Y^{-1}(B)) = 0$ for $\omega \notin Y^{-1}(B)$ and $P(\omega \mid Y^{-1}(B)) = P(\omega)/PY^{-1}(B))$ for $\omega \in Y^{-1}(B)$.

$E(X \mid Y \in B)$ is just the expectation of the random variable X restricted to the instances ω for which $Y(\omega)$ takes a value in B. The individual probabilities of these instances are just normalized by $P(Y \in B)$ so that their sum is 1 while they maintain the same relative probabilities with respect to each other as before the conditioning.

Now it is well known (see, for example, [11, p. 144]) that if $Y^{-1}(B)$ and $Z^{-1}(D)$ are disjoint and exhaustive of Ω, that is, $Y^{-1}(B) \wedge Z^{-1}(D) = \Phi$ and $Y^{-1}(B) \vee Z^{-1}(D) = \Omega$ and if X, Y, W and Z are random variables on Ω, then

$$\begin{aligned} &E((X \mid Y \in B) + (W \mid Z \in D)) \\ &\quad = E(X \mid Y \in B) P(Y \in B) + E(W \mid Z \in D) P(Z \in B). \qquad (26) \end{aligned}$$

That is, the expectation of the sum of the conditional random variables is the sum of the conditional expectations weighted by the probabilities of the associated conditions. With the extended definitions of operations on random variables this result can be generalized to allow $Y^{-1}(B)$ and $Z^{-1}(D)$ to be arbitrary events that may overlap and also may not be exhaustive of Ω.

First we extend the result to disjoint events $Y^{-1}(B)$ and $Z^{-1}(D)$ that do not necessarily exhaust Ω.

LEMMA 1. *If $Y^{-1}(B)$ and $Z^{-1}(D)$ are disjoint events of Ω, and if X, Y, W and Z are random variables on Ω, then*

$$E((X \mid Y \in B) + (W \mid Z \in D))$$

$$= E(X \mid Y \in B)P(Y \in B \mid Y \in B \vee Z \in D) \tag{27}$$
$$+ E(W \mid Z \in D)P(Z \in B \mid Y \in B \vee Z \in D)).$$

Proof. This result follows by using a new probability measure on just the part of Ω inside $(Y \in B \vee Z \in D)$. So let Q be the probability measure defined by $Q(A) = P(A \mid Y \in B \vee Z \in D)$ for any event A in Ω. That Q is a probability measure on $(Y \in B \vee Z \in D)$ is easy to show since it is non-negative, $Q(Y \in B \vee Z \in D) = P(Y \in B \vee Z \in D \mid Y \in B \vee Z \in D) = 1$, and finally, if A and C are disjoint events in Ω, then $Q(A \vee C) = P(A \vee C \mid Y \in B \vee Z \in D) = P(A \vee C)/P(Y \in B \vee Z \in D) = (P(A) + P(C))/P(Y \in B \vee Z \in D) = Q(A) + Q(B)$. In addition, the conditional expectation $E_Q(X \mid Y \in B)$ with respect to Q of an arbitrary random variable X given arbitrary $(Y \in B)$ equals the conditional expectation $E_P(X \mid Y \in B)$ with respect to P because

$$\begin{aligned} E_Q(X \mid Y \in B) &= \sum_{\omega\in\Omega} X(\omega)Q(\omega \wedge (Y \in B) \mid Y \in B) \\ &= \sum_{\omega\in\Omega} X(\omega)Q(\omega \wedge (Y \in B))/Q(Y \in B) \\ &= \sum_{\omega\in\Omega} X(\omega)P(\omega \wedge (Y \in B))/P(Y \in B) \\ &= E_P(X \mid Y \in B). \end{aligned} \tag{28}$$

So now computing

$$\begin{aligned} &E_P((X \mid Y \in B) + (W \mid Z \in D)) \\ &\quad = E_Q((X \mid Y \in B) + (W \mid Z \in D)) \\ &\quad = E(X \mid Y \in B)Q(Y \in B) + E(W \mid Z \in D)Q(Z \in D) \\ &\quad = E(X \mid Y \in B)P(Y \in B \mid Y \in B \vee Z \in D) \\ &\qquad + E(W \mid Z \in D)P(Z \in D \mid Y \in B \vee Z \in D). \end{aligned} \tag{29}$$

That completes the proof of Lemma 1. □

THEOREM 1. *If X, Y, W and Z are real-valued random variables and B and D are arbitrary Borel subsets of real numbers, then*

$$\begin{aligned} &E((X \mid Y \in B) + (W \mid Z \in D)) \\ &\quad = E(X \mid Y \in B)P(Y \in B \mid Y \in B \vee Z \in D) \\ &\qquad + E(W \mid Z \in D)P(Z \in D \mid Y \in B \vee Z \in D). \end{aligned} \tag{30}$$

Proof. Let $K = Y^{-1}(B) = \{\omega \in \Omega : Y(\omega) \in B\}$ and $L = Z^{-1}(D)$. So using the definition of extended summation for conditional random variables,

$$\begin{aligned} & E((X \mid Y \in B) + (W \mid Z \in D)) \\ & \quad = E((X \mid K) + (W \mid L)) = E(X + W \mid K \vee L) \\ (31) \quad & \quad = E((X \mid KL') + (X + W \mid KL) + (W \mid K'L)), \end{aligned}$$

where juxtaposition has again replaced conjunction ($\wedge$) to shorten notation.

Since KL', KL and $K'L$ are disjoint, according to Lemma 1, we can continue with

$$\begin{aligned} & E((X \mid KL') + (X + W \mid KL) + (W \mid K'L)) \\ & \quad = E(X \mid KL')P(KL' \mid K \vee L) \\ & \qquad + E(X + W \mid KL)P(KL \mid K \vee L) \\ (32) \quad & \qquad + E(W \mid K'L)P(K'L \mid K \vee L) \\ & \quad = E(X \mid KL')P(KL' \mid K \vee L) \\ & \qquad + E(X \mid KL)P(KL \mid K \vee L) \\ & \qquad + E(W \mid KL)P(KL \mid K \vee L)) \\ (33) \quad & \qquad + E(W \mid K'L)P(K'L \mid K \vee L) \\ & \quad = E((X \mid KL')P(KL' \mid K)P(K \mid K \vee L) \\ & \qquad + E(X \mid KL)P(KL \mid K)P(K \mid K \vee L) \\ & \qquad + E(W \mid KL)P(KL \mid L)P(L \mid K \vee L) \\ (34) \quad & \qquad + E(W \mid K'L)P(K'L \mid L)P(L \mid K \vee L) \\ & \quad = [E((X \mid KL')P(KL' \mid K) \\ & \qquad + E(X \mid KL)P(KL \mid K)]P(K \mid K \vee L) \\ & \qquad + [E(W \mid KL)P(KL \mid L) \\ (35) \quad & \qquad + E(W \mid K'L)P(K'L \mid L)]P(L \mid K \vee L) \\ & \quad = E((X \mid KL') + (X \mid KL))P(K \mid K \vee L) \\ (36) \quad & \qquad + E((W \mid KL) + (W \mid K'L))P(L \mid K \vee L), \end{aligned}$$

using Lemma 1 in reverse. So

$$\begin{aligned} & E((X \mid K) + (W \mid L)) \\ (37) \quad & \quad = E(X \mid K)P(K \mid K \vee L) + E(W \mid L)P(L \mid K \vee L). \end{aligned}$$

That is,

$$E((X \mid Y \in B) + (W \mid Z \in D))$$

$$= E(X \mid Y \in B)P(Y \in B \mid Y \in B \vee Z \in D)$$
$$+ E(W \mid Z \in D))P(Z \in D \mid Y \in B \vee Z \in D). \tag{38}$$

That completes the proof of Theorem 1. □

THEOREM 2. *If X, Y, W and Z are real-valued random variables and B and D are arbitrary Borel subsets of real numbers, then the expectation of the product of the conditional random variables $(X \mid Y \in B)$ and $(W \mid Z \in D)$ is given by*

$$\begin{aligned}
&E((X \mid Y \in B) * (W \mid Z \in D)) \\
&\quad = E(X \mid Y \in B \wedge Z \notin D)P(Y \in B \wedge Z \notin D \mid \\
&\qquad Y \in B \vee Z \in D) \\
&\qquad + E(X * W \mid Y \in B \wedge Z \in D)P(Y \in B \wedge Z \in D \mid \\
&\qquad Y \in B \vee Z \in D) \\
&\qquad + E(Z \mid Y \notin B \wedge Z \in D)PY \notin B \wedge Z \in D \mid \\
&\qquad Y \in B \vee Z \in D).
\end{aligned} \tag{39}$$

Proof. By the extended definition of products,

$$(X \mid Y \in B) * (W \mid Z \in D)$$
$$= \begin{cases} X & \text{if } Y \in B \text{ and } Z \notin D, \\ X * W & \text{if } Y \in B \text{ and } Z \in D, \\ W & \text{if } Y \notin B \text{ and } Z \in D, \\ \text{Undefined} & \text{if } Y \notin B \text{ and } Z \notin D, \end{cases} \tag{40}$$

where the domain of the product random variable has been broken into disjoint events. Then by the definition of the conditional expectation (Equation (25)), the expectation of this product random variable, $E((X \mid Y \in B) * (W \mid Z \in D))$, is immediately expressed by Equation (39). This completes the proof of Theorem 2. □

With a kind of independence, Equation (39) can be somewhat simplified. Recall that two random variables, X and Z, are independent if $P(X \in A \text{ and } Z \in C) = P(X \in A)P(Z \in C)$ for any events $X \in A$ and $Z \in C$. Knowing the value taken by one variable does not change the probability of the other variable taking its values.

DEFINITION 1 (Independence of Random Variables). Two random variables X and W are independent if they are independent on each common domain. That is, X and W are independent if for any event H for which both X and W are defined, X is conditionally independent of W given H. That is $P(X \in A \wedge W \in C \mid H) = P(X \in A \mid H)P(W \in C \mid H)$.

COROLLARY TO THEOREM 2. *If X and W are independent random variables then under the hypothesis of Theorem 2,*

$$\begin{aligned}
&E((X \mid Y \in B) * (W \mid Z \in D)) \\
&\quad = E(X \mid Y \in B \wedge Z \notin D)P(Y \in B \wedge Z \notin D \mid \\
&\qquad Y \in B \vee Z \in D) \\
&\qquad + E(X \mid Y \in B \wedge Z \in D)E(W \mid \\
&\qquad Y \in B \wedge Z \in D)P(Y \in B \wedge Z \in D \mid \\
&\qquad Y \in B \vee Z \in D) \\
&\qquad + E(W \mid Y \notin B \wedge Z \in D)P(|Y \notin B \wedge Z \in D \mid \\
&\qquad Y \in B \vee Z \in D).
\end{aligned} \tag{41}$$

Proof. It is well known that the expectation of a product of independent random variables is the product of the expectations. Therefore $E(X * W \mid Y \in B \wedge Z \in D) = E(X \mid Y \in B \wedge Z \in D)E(W \mid Y \in B \wedge Z \in D)$, and the result follows by substitution into Equation (39). □

EXAMPLE 1. Consider a work force consisting of workers $i = 1, 2, \ldots, n$ with variable work output levels $W_1, W_2, \ldots, W_n$ and work shifts $s_1, s_2, \ldots, s_n$ respectively spanning the 24 hour day. To formulate the problem in terms of random variables, let $s_i(\omega) = 1$ if time $\omega \in s_i$ and 0 otherwise. Then the work level at time ω of worker i is $(W_i \mid s_i(\omega) = 1)$. The sum of work output of all workers is $\sum_i (W_i \mid s_i(\omega) = 1)$, and the average or expected work level over the day is $E(\sum_i (W_i \mid s_i(\omega) = 1)) = \sum_i E(W_i \mid S_i(\omega) = 1)P(s_i(\omega) = 1))$.

EXAMPLE 2. Let B and C be the surveillance regions of two radars, R1 and R2, and suppose $X(\omega)$ is the error rate of missed detections by R1 at any place $\omega \in B$, and $W(\omega)$ is the error rate by R2 at any place $\omega \in C$. X and W are undefined outside their respective domains B and C. Then using the definition of extended product, and assuming independence of detections by R1 and R2, $(X * W)(\omega) = X(\omega)W(\omega)$ is the combined error rate of missed detections by both radars over $(B \cup C)$. This combined error rate is X on $B \cap C'$, $X * W$ on $B \cap C$, and W on $B' \cap C$.

Now suppose in addition that the detection rate of radar R1 is greatly affected by fog F while interrogation radar R2 is most affected by the density D of communication on interrogation frequencies. Measuring fog as "heavy (h), medium (m), or none (n)" and communication traffic density on a scale from 1 to 3, the error rate over $(B \cup C)$ under conditions of heavy fog and communication density 2 is $((X * W) \mid (F = h) \wedge (D = 2) \wedge (B \cup C)) = ((X * W) \mid (F = h)(D = 2)(B \cup C))$. So the expected combined error rate of the two radars given heavy fog and medium (2) communication density is $E((X * W) \mid (F = h)(D = 2)(B \cup C))$.

Now by the product definition

$$\begin{aligned} &((X * W) \mid (F = h)(D = 2)(B \cup C)) \\ &\quad = (X \mid (F = h)(D = 2)BC') \vee (X * W \mid \\ &\qquad (F = h)(D = 2)BC) \vee (W \mid (F = h)(D = 2)B'C). \end{aligned} \tag{42}$$

Since the detection errors for the two radars are assumed independent, the last equation simplifies to

$$\begin{aligned} &((X * W) \mid (F = h)(D = 2)(B \cup C)) \\ &\quad = (X \mid (F = h)BC') \vee (X * W \mid \\ &\qquad (F = h)(D = 2)BC) \vee (W \mid (D = 2)B'C). \end{aligned} \tag{43}$$

Let $G = (F = h)(D = 2)(B \cup C)$. Then in terms of the average error rates of the individual radars, the average combined error rate given heavy fog and medium (2) communication density over the combined surveillance region $(B \cup C)$ is

$$\begin{aligned} &E((X * W) \mid (F = h)(D = 2)(B \cup C)) \\ &\quad = E(X \mid (F = h)BC')P((F = h)BC') \mid G) \\ &\qquad + E(X * W \mid (F = h)(D = 2)BC)P((F = h)(D = 2)BC \mid \\ &\qquad G) + E(W \mid (D = 2)(BC'))P(D = 2)B'C \mid G) \end{aligned} \tag{44}$$

$$\begin{aligned} &\quad = E(X \mid (F = h)(BC'))P((F = h)(BC') \mid G) \\ &\qquad + E(X \mid (F = h)(BC))E(W \mid \\ &\qquad (D = 2)(BC))P((F = h)(D = 2)(BC) \mid G) \\ &\qquad + E(W \mid (D = 2)(B'C))P((D = 2)(B'C) \mid G) \end{aligned} \tag{45}$$

using conditional independence again to split the expectation of $X * W$ and to simplify the conditions.

For simplicity, assume that $B \cup C$ is the whole universe and that fog is heavy $(F = h)$ everywhere and communication density is medium $(D = 2)$ everywhere in $B \cup C$. So $G =$ the whole universe Ω, and $P((F = h)BC') \mid G) = P(BC')$. Similarly, $P((F = h)(D = 2)BC \mid G) = P(BC)$ and $P((D = 2)(B'C) \mid G) = P(B'C)$. Thus

$$E((X * W) \mid (F = h)(D = 2)(B \cup C))$$

$$
\begin{aligned}
&= E(X \mid (F = h)(BC'))P(BC') \\
&\quad + E(X \mid (F = h)(BC))E(W \mid (D = 2)(BC))P(BC) \\
&\quad + E(W \mid (D = 2)(B'C))P(B'C).
\end{aligned} \tag{46}
$$

If the error rate of radar R1 is 0.04 in heavy fog ($F = h$) and the error rate of R2 is 0.02 in medium communication density ($D = 2$), the combined error rate under the conditions is

$$
\begin{aligned}
&E((X * W) \mid (F = h)(D = 2)(B \cup C)) \\
&\quad = (0.04)P(BC') + (0.04)(0.02)P(BC) + (0.02)P(B'C).
\end{aligned} \tag{47}
$$

Note that the error rates are multiplied in the common surveillance region BC where the combined error rate is just 0.0008.

11. Summary

Extended definitions of function addition and other operations have been applied to conditional propositions and conditional events, and to conditional random variables. This allows direct manipulation of conditional events and of conditional random variables without resort to a probability or density function. General formulas for the expectation of the sum, and of the product, of two conditional random variables have been determined. Finally two examples illustrate the use of these formulas in practical situations.

References

1. Adams, E. W.: Probability and the logic of conditionals, in P. Suppes and J. Hintikka (eds.), *Aspects of Inductive Logic*, North-Holland, Amsterdam, 1966, pp. 265–316.
2. Calabrese, P. G.: An algebraic synthesis of the foundations of logic and probability, *Inform. Sci.* **42** (1987), 187–237.
3. Calabrese, P. G.: Reasoning with uncertainty using conditional logic and probability, in *Proc. First International Symposium on Uncertainty Modeling and Analysis*, IEEE Computer Society, 1990, pp. 682–688.
4. Calabrese, P. G.: Deduction and inference using conditional logic and probability, in I. R. Goodman, M. M. Gupta, H. T. Nguyen and G. S. Rogers (eds.), *Conditional Logic in Expert Systems*, North-Holland, Amsterdam, 1991, pp. 71–100.
5. Calabrese, P. G.: A theory of conditional information with applications, Special Issue on Conditional Event Algebra, *IEEE Trans. Systems Man Cybernet.* **24**(12) (1994), 1676–1684.

6. Calabrese, P. G.: An extension of the fundamental theorem of Boolean algebra to conditional propositions, Part I (1–18) of "Conditional event algebras and conditional probability logics" by P. G. Calabrese and I. R. Goodman, in R. Scozzafava (ed.), *Probabilistic Methods in Expert Systems, Proceedings of the International Workshop*, Rome, Italy, Societá Italiana di Statistica, 1996, pp. 1–35.
7. Calabrese, P. G.: Deduction with uncertain conditionals, *Inform. Sci.*, in press.
8. De Finetti, B.: La logique de la probabilite, in *Actes du Congres International de Philosophie Scientifique IV*, Hermann, Paris, 1936, pp. 1–8.
9. Dubois, D. and Prade, H.: Conditional objects as nonmonotonic consequence relationships, *IEEE Trans. Systems Man Cybernet.* **24**(12) (1994), 1724–1740.
10. Kolmogorov, A. N.: *Foundations of the Theory of Probability*, Chelsea, 1956 (1st edition: *Grundbegriffe der Wahrscheinlichkeits-rechnung*, Berlin, 1933).
11. Papoulis, A.: *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1965.
12. Suppes, P. and Zanotti, M.: *Z. Wahrsch. verw. Gebiete* **60** (1982), 163–169. (Also in [13].)
13. Suppes, P. and Zanotti, M.: *Foundations of Probability with Applications*, Selected Papers 1974–1995, Cambridge Univ. Press, 1996.

Joint & National Systems Division (Code 2737),
Space & Naval Warfare Systems Center San Diego,
CA 92152-5001, USA
(e-mail: calabres@spawar.navy.mil)