Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 272 (2015) 89–101
FUZZY sets and systems
www.elsevier.com/locate/fss

Conditional information using copulas with an application to decision making

Ronald R. Yager

Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, United States

Received 24 April 2014; received in revised form 26 January 2015; accepted 28 January 2015

Available online 3 February 2015

Abstract

We focus on the task of calculating conditional probabilities of the form $\text{Prob}(U \leq x|V \leq y)$. We point out that in this case the relevant probabilities, $\text{Prob}(U \leq x)$ and $\text{Prob}(V \leq y)$, have the nature of a cumulative distribution. This enables us to use the Sklar theorem to directly calculate the required joint probability as a simple binary aggregation of these marginals using a copula. Here the choice of copula reflects the type of correlation between U and V. We study in considerable detail the effects of using different copulas. We also show that this enables us to simply and directly calculate the probability that $U = x$ conditioned on the knowledge of the $\text{Prob}(V \leq y)$. We use this result to aid in decision-making where we compare alternative's expected payoffs based on the conditioned probabilities.

Keywords: Conditional probabilities; Copula; Uncertain decision making

1. Introduction

Conditional probability provides a very useful tool for modifying "a priori" probability information about a variable of interest based upon probability information of a related variable. This can be particularly valuable in decision-making tasks where the conditioned probability provides more informed knowledge about the environment. Implicit in the formulation of the conditional probability is the requirement of having information about the joint probabilities of the conditioned and conditioning variables. A beautiful result that can help us, at times, in this task is contained in the Sklar theorem [1]. This theorem provides a direct way of obtaining the joint cumulative distribution function from the two marginal cumulative distribution functions by a simple binary aggregation of these marginals using a copula [2–7], which is a kind of "anding" operator [8]. Here the choice of the copula is a reflection of the type of correlation between the conditioned and conditioning variables. Recalling that a cumulative distribution function expresses the probability that a variable is less then or equal a value, $\text{Prob}(U \leq x)$, we focus on the calculation of conditional probabilities of the form $\text{Prob}(U \leq x|V \leq y)$. We see here that $\text{Prob}(U \leq x|V \leq y) = \frac{\text{Prob}(U \leq x, V \leq y)}{\text{Prob}(V \leq y)}$. Since $\text{Prob}(U \leq x, V \leq y)$ is essentially the determination of a joint cumulative distribution from marginal cumu-

E-mail address: yager@panix.com.

http://dx.doi.org/10.1016/j.fss.2015.01.018
0165-0114/

lative distributions we are able to take advantage of the Sklar theorem to calculate these conditional probabilities. We first look at this under the use of different copulas. A further useful result is that in the discrete environment since $\text{Prob}(U = x_i | V \le y) = \text{Prob}(U \le x_i | V \le y) - \text{Prob}(U \le x_{i-1} | V \le y)$ then we can directly calculate the probability that $U = x_i$ conditioned on the knowledge of the $\text{Prob}(V \le y)$. We illustrate the use of this in the case of decision-making under uncertainty where we can compare alternative's expected payoffs based on the conditioned probabilities.

2. Conditional probabilities and cumulative distribution functions

Assume U is a random variable on the ordered space $X = \{x_1, \dots, x_n\}$ with $x_i \le x_{i+1}$ and where p_i is the probability that $U = x_i$. The cumulative distribution function (CDF), $F : X \to [0, 1]$ is defined so $F_1(x_j) = \text{Prob}(U \le x_j) = \sum_{i=1}^{j} p_i$.

Another important concept in probability theory is conditional probability. We recall $\text{Prob}(A|B) = \frac{\text{Prob}(A \cap B)}{\text{Prob}(B)}$ and $\text{Prob}(B) > 0$. We emphasize that if A and B are subsets of the same probability space X then $A \cap B$ are the elements that are in both A and B. If A and B are defined with respect to different spaces, X and Y respectively, then $A \cap B = A \times B$, the Cartesian product.

Assume U and V are two random variables taking their values in the ordered spaces X and Y respectively. Here $x_i \ge x_j$ if $i > j$ and $y_i \ge y_k$ if $i > k$. We can associate with each of a CDF, F_U and F_V respectively, so that $F_U(x_j) = \text{Prob}(U \le x_j) = \sum_{k=1}^{j} \text{Prob}(U = x_k)$ and $F_V(y_i) = \text{Prob}(V \le y_i) = \sum_{k=1}^{i} \text{Prob}(V = y_k)$ when $\text{Prob}(U = x_j) > 0$ and $\text{Prob}(V = y_i) > 0$. We can introduce the idea of a joint CDF, $F_{U,V}(x_j, y_i) = \text{Prob}(U \le x_j, V \le y_i)$ where $\text{Prob}(U \le x_j, V \le y_i) = \sum_{k=1}^{j} \sum_{r=1}^{i} \text{Prob}(U = x_k, V = y_r)$ where $\text{Prob}(U = x_k, V = y_r)$ is the probability of the joint outcome with $U = x_k$ and $V = y_r$.

We point out that $\{U \le x_j\}$ is an event thus $\text{Prob}(U \le x_j)$ is simply the probability of some event. The same is true of $\text{Prob}(V \le y_i)$. Let us denote $A = \{U \le x_j\}$ and $B = \{V \le y_i\}$. We can use these in our formulation for conditional probability

$$\text{Prob}(A|B) = \frac{\text{Prob}(A \cap B)}{P(B)} = \frac{\text{Prob}(U \le x_j, V \le y_i)}{\text{Prob}(V \le y_i)} = \text{Prob}(U \le x_j | V \le y_i) = \frac{F_{U,V}(x_j, y_i)}{F_V(y_i)}.$$

We note that $\text{Prob}(U \le x_j | V \le y_i) - \text{Prob}(U \le x_{j-1} | V \le y_i) = \frac{F_{U,V}(x_j, y_i)}{F_V(y_i)} - \frac{F_{U,V}(x_{j-1}, y_i)}{F_V(y_i)}$ and hence $\frac{F_{U,V}(x_j, y_i)}{F_V(y_i)} - \frac{F_{U,V}(x_{i-1}, y_i)}{F_V(y_i)} = \frac{F_{U,V}(x_j, y_i) - F_{U,V}(x_{j-1}, y_i)}{F_V(y_i)}$. We further observe that

$$F_{U,V}(x_j, y_i) = \sum_{k=1}^{j} \sum_{r=1}^{i} \text{Prob}(U = x_k, V = y_r) \quad \text{and} \quad F_{U,V}(x_{j-1}, y_i) = \sum_{kl=1}^{j-1} \sum_{r=1}^{i} \text{Prob}(U = x_k, V = y_r).$$

Using this we get that $F_{U,V}(x_j, y_i) - F_{U,V}(x_j - 1, y_i) = \sum_{r=1}^{i} \text{Prob}(U = x_j, V = y_r)$ but this is simply the probability of the joint event, $\{U = x_j\} \times \{V \le y_i\}$. Denoting $A = \{U = x_j\}$ and $B = \{V \le y_i\}$ we get $\frac{F_{U,V}(x_j, y_i)}{F_V(y_i)} - \frac{F_{U,V}(x_{j-1}, y_i)}{F_V(y_i)} = \frac{\text{Prob}(A \times B)}{\text{Prob}(B)}$ the conditional probability $\text{Prob}(A|B)$ where $A = \{U = x_j\}$ and $B = \{V \le y_i\}$. Hence we see that

$$\text{Prob}(U = x_j | V \le y_i) = \frac{F_{U,V}(x_j, y_i)}{F_V(y_i)} - \frac{F_{U,V}(x_{j-1}, y_i)}{F_V(y_i)}$$
$$\text{Prob}(U = x_j | V \le y_i) = \text{Prob}(U \le x_j | V \le y_i) - \text{Prob}(U \le x_{j-1} | V \le y_i)$$

We note that if D is any subset associated with the space X then

$$\text{Prob}(U \in D | V \le y_i) = \sum_{x_j \in D} \text{Prob}(U = x_j | V \le y_i)$$

More generally we can show that for any subset B on the space of V that

$$\text{Prob}(U = x_j | V \in B) = \text{Prob}\big(U \in D \cup \{x_j\} | V \in B\big) - \text{Prob}(U \in D | V \in B)$$

where D is any subset on the space X which does not contain x_j.