STA 623 – Fall 2011 – Dr. Charnigo

Section 1.3: Conditional Probability and Independence

Definition of conditional probability. Let A and B be events that belong to the sigma field. (Hereafter all events mentioned will belong to the sigma field unless explicitly stated otherwise.) If $P(B) > 0$, then we define the conditional probability of A given B as

$$P(A|B) := P(A \cap B)/P(B).$$

In this context, we sometimes refer to $P(A)$ as an unconditional probability. Intuitively, $P(A|B)$ is an updated version of $P(A)$ given the knowledge that event B has occurred.

Example (definition of conditional probability). Let A denote the event that a randomly selected person smokes, and let B denote the event that a randomly selected person develops lung cancer. Suppose that $P(B) = 0.02$, $P(A) = 0.25$, and $P(B|A) = 0.05$. Then $P(A \cap B) =$ ________, so that $P(A|B) =$ ________. Moreover, $P(A^c \cap B) =$ ________, so that $P(A^c|B) =$ ________. Is the last answer intuitively obvious?

Useful results for conditional probabilities. Assuming that the axioms for unconditional probabilities are satisfied and that $P(D) > 0$, we have the following useful results for conditional probabilities.

1. $P(A|D) \geq 0$.
2. $P(S|D) = 1$.
3. $P(\cup_{i=1}^{\infty} A_i|D) = \sum_{i=1}^{\infty} P(A_i|D)$ whenever $A_i \cap A_j = \emptyset$ for $i \neq j$.
4. $P(\emptyset|D) = 0$.

5. $P(A|D) \leq 1$.
6. $P(A^c|D) = 1 - P(A|D)$.
7. $P(B \cap A^c|D) = P(B|D) - P(B \cap A|D)$.
8. $P(A \cup B|D) = P(A|D) + P(B|D) - P(A \cap B|D)$.
9. If $A \subset B$, then $P(A|D) \leq P(B|D)$.
10. $P(\cup_{i=1}^{\infty} C_i|D) \leq \sum_{i=1}^{\infty} P(C_i|D)$.
11. If $C_1, C_2, \ldots$ is a partition, then $P(A|D) = \sum_{i=1}^{\infty} P(A \cap C_i|D)$.

Example (useful results for conditional probabilities). To verify result 3, we write

$$P(\cup_{i=1}^{\infty} A_i|D) = P(\cup_{i=1}^{\infty}(A_i \cap D))/P(D) = \qquad\qquad = \sum_{i=1}^{\infty} P(A_i|D).$$

The justification for the second equality is that, if $A_i \cap A_j = \emptyset$, then

$$(A_i \cap D) \cap (A_j \cap D) = (A_i \cap A_j) \cap D = \emptyset \cap D = \emptyset.$$

To verify result 9, we note that $A \cap D \subset B \cap D$. (How do we know this?) Then $P(A \cap D) \leq P(B \cap D)$, whence

$$P(A|D) = P(A \cap D)/P(D) \leq P(B \cap D)/P(D) = P(B|D).$$

Iterating conditional probabilities. Suppose that $P(B \cap C) > 0$. Then

$$P(A \cap B \cap C) = P(A|B \cap C)P(B \cap C) = P(A|B \cap C)P(B|C)P(C).$$

Moreover, since the left side is also equal to $P(A \cap B|C)P(C)$, we see that

$$P(A \cap B|C) = P(A|B \cap C)P(B|C).$$

(How do we know that $P(C) > 0$?) The above formulas may be useful in situations when the conditional probabilities $P(A|B \cap C)$ and $P(B|C)$ are intuitively obvious while $P(A \cap B|C)$ and $P(A \cap B \cap C)$ are less readily apparent.

Example (iterating conditional probabilities). I am dealt three cards (without replacement) from a well-shuffled standard 52-card deck. Let C be the event that the suit of the first card is diamonds, B be the event that the suit of the second card is diamonds, and A be the event that the suit of the third card is diamonds. Then $P(C) = 13/52$, $P(B|C) =$, and $P(A|B \cap C) =$. So $P(A \cap B|C) =$ and $P(A \cap B \cap C) =$.

Bayes' Theorem. Let $A_1, A_2, \ldots$ be a partition of S such that $P(A_i) > 0$ for $i \in \{1, 2, \ldots\}$. Then for any event B and any $i \in \{1, 2, \ldots\}$ we have

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}.$$

A similar result also holds for a finite partition $A_1, A_2, \ldots, A_k$,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{k} P(B|A_j)P(A_j)}.$$

In particular, with $k = 2$ we have (upon a minor change in notation)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}.$$

This last version of Bayes' Theorem provides a recipe for computing $P(A|B)$ if we have available $P(B|A)$ (and two other probabilities — why two rather than three?).

Example (Bayes' Theorem). Bayes' Theorem is obvious from the definition of conditional probability if we can verify that the denominator is $P(B)$. (Equality of the denominator to $P(B)$ is sometimes called the Law of Total Probability.) Suggestions?

Again, let A denote the event that a randomly selected person smokes, and let B denote the event that a randomly selected person develops lung cancer. As before, suppose that $P(A) = 0.25$ and $P(B|A) = 0.05$. But now suppose that we are not given $P(B)$. Rather, suppose that we are told that 99% of non-smokers do not develop lung cancer. How can we find $P(A|B)$ with just the information provided here?

Independence of two events. Suppose that A and B are events such that $0 < P(A) < 1$, $0 < P(B) < 1$, and $P(A|B) = P(A)$. Thus, knowing that B has occurred does not lead us to revise the probability that A will occur. In this case, multiplying both sides of $P(A|B) = P(A)$ by $P(B)$ yields

$$P(A \cap B) = P(A|B)P(B) = P(A)P(B).$$

Moreover, since $P(A \cap B)$ can also be expressed as $P(B|A)P(A)$, we see that $P(B|A) = P(B)$. If any one of these three equivalent conditions holds — $P(A|B) = P(A) = P(A|B^c)$, $P(A \cap B) = P(A)P(B)$, or $P(B|A) = P(B) = P(B|A^c)$ — we say that A and B are independent. (How are we able to go from $P(A|B) = P(A)$ to $P(A|B) = P(A) = P(A|B^c)$ and from $P(B|A) = P(B)$ to $P(B|A) = P(B) = P(B|A^c)$?)

If no restrictions on $P(A)$ or $P(B)$ are made, so we are not sure whether $P(A|B)$ and $P(B|A)$ are defined, then we characterize independence by the equation $P(A \cap B) = P(A)P(B)$.

If A and B are independent, then so are A^c and B^c, A^c and B, and A and B^c.

Example (independence of two events). Let A and B be two events. If $P(A) = 0$, then A and B are independent. (How do we know this?) If $P(B) = 1$, then A and B are independent. (How do we know this?)

To verify that independence of A and B implies independence of A^c and B, we write

$$P(A^c \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = P(B)P(A^c).$$

We can similarly prove that independence of A and B implies independence of A^c and B^c and of A and B^c.

Suppose once more that I am dealt three cards (without replacement) from a well-shuffled standard 52-card deck. Let C be the event that the suit of the first card is diamonds, B be the event that the suit of the second card is diamonds, and A be the event that the suit of the third card is diamonds. Are C and B independent?

Independence of three or more events. Let $A_1, A_2, \ldots, A_n$ be events such that for any subcollection $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ we have

$$P(\cap_{j=1}^k A_{i_j}) = \prod_{j=1}^k P(A_{i_j}).$$

Then we say that $A_1, A_2, \ldots, A_n$ are independent. Independence of n events thus entails not only

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n)$$

but also

$$P(A_1 \cap A_2) = P(A_1)P(A_2),$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3),$$
$$P(A_4 \cap A_{18}) = P(A_4)P(A_{18}),$$

and many other equalities.