# The Average Rule and Bayes' Rule

## The formulas

1. **Partitions:** A collection of sets $B_1, B_2, \dots, B_n$ is said to **partition** the sample space $S$ if the sets (i) are mutually disjoint and (ii) have as union the entire sample space. In particular, any set $B$, together with its complement $B^c$ forms a partition of $S$. More generally, given a set $B$, $B_1, B_2, \dots, B_n$ are said to partition a set $B$ if the sets (i) are mutually disjoint and (ii) their union is $B$.

2. **Average and Bayes' Rules, special case:** For any events $A$ and $B$,

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) \quad \text{(Average Rule)} \tag{1}$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)} \quad \text{(Bayes' Rule)} \tag{2}$$

3. **Average and Bayes' Rules, general case:** Suppose $B_1, B_2, \dots, B_n$ partition the sample space. Then for each $i = 1, 2, \dots, n$ and any set $A$,

$$P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n) \quad \text{(Average Rule)} \tag{3}$$

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)} \quad \text{(Bayes' Rule)} \tag{4}$$

## Notes

1. **Technical note:** The probabilities $P(A|B_i)$ are only defined if $P(B_i) > 0$; if $P(B_i) = 0$, the corresponding term $P(A|B_i)P(B_i)$ in the above formulas should be interpreted as 0 (i.e., it doesn't affect the formulas).

2. **Memorizing Bayes' Rule:** The Average Rule says that the expression appearing in the denominator in Bayes' Rule is equal to $P(A)$. Thus, Bayes' rule could have also been stated in the form

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)}. \tag{5}$$

   You could get by memorizing the average rule (3) along with the above simpler version of Bayes' Rule (5). However, I recommend memorizing Bayes' Rule in the full form (4), since that is the form that you normally need in applications.

3. **General versus special case of Bayes' Rule:** Many, but not all, applications of Bayes' Rule involve only the special case when the simpler formula (2) can be used. However, for more general problems, one does need the more complicated formula (4). I recommend to memorize both formulas.

## Interpretations and Applications

1. **Interpretation of the Average Rule:** A partition $B_1, \dots, B_n$ of $S$ can be interpreted as a set of mutually exclusive conditions that exhaust all possibilities. With this interpretation, the average rule says that the (overall) probability of $A$ is equal to a weighted average of all the conditional probabilities, $P(A|B_i)$, of $A$ with respect to $B_i$, with the weigths being the probabilities with which these conditions occur.

2. **Interpretation of Bayes' Rule:** Bayes' Rule can be interpreted in terms of *prior* and *posterior* probabilities. The prior probabilities are $P(B_i)$, i.e., the (ordinary) probability that the event $B_i$ occurs. Bayes' Rule shows how these probabilities change if we know that event $A$ has occurred; namely it gives a formula for $P(B_i|A)$, the conditional probability that $B_i$ occurs given that $A$ has occurred. The latter probabilities are called posterior probabilities. (The terms "prior" and "posterior" come from Latin and mean "before" and "after".)

3. **Recognizing Bayes' Rule problems:** Bayes' Rule is a formula for reversing the order in conditional probabilities. Many (but not all) conditional probability problems are of this type. If the probability sought in the problem is a conditional probability and the same conditional probability, but with the order of events reversed, is given (or can easily be deduced from the given information), the problem is likely a Bayes' Rule problem. Example: In the blood test problem, the probability sought is that of someone having the disease given that he/she tests positive, whereas the reverse conditional probability, that someone tests positive given that he/she has the disease, is given.

4. **Recognizing conditional probabilities:** Conditional probabilities are often indicated by words/phrases like "given that", or "if", or by words implying a *subpopulation*. Here are some examples of statements (mostly taken from actuarial exam problems) that refer to a conditional probability, along with their translation into mathematical language. The "give-away" words that indicate that a conditional probability is involved are set in italics.

   - "*Among* independents, McCain leads Bush 54% to 36%."
     Translation: $P(M|I) = 0.54$ and $P(B|I) = 0.36$, where $M$="McCain voter", $B$="Bush voter", $I$="Independent".
   - "5 percent *of those* having the disease test negative."
     Translation: $P(N|D) = 0.05$, where $N$="test negative", $D$="has disease".
   - "*For each smoker*, the probability of dying during the year is 0.05"
     Translation: $P(D|S) = 0.05$, where $D$="dies during the year", $S$="is a smoker".
   - "A blood test indicates the presence of a disease 95% of the time the disease is actually present."
     Translation: $P(P|D) = 0.95$, where $P$="tests positive", $D$="has disease".