General Probability, III: Bayes' Rule

Bayes' Rule

1. **Partitions:** A collection of sets $B_1, B_2, \ldots, B_n$ is said to **partition** the sample space if the sets (i) are mutually disjoint and (ii) have as union the entire sample space. A simple example of a partition is given by a set B, together with its complement B'.

2. **Total Probability Rule (Average Rule):** If $B_1, B_2, \ldots, B_n$ partition the sample space, then for any set A,
$$P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n). \tag{1}$$

3. **Bayes' Rule, general case:** If $B_1, B_2, \ldots, B_n$ partition the sample space, then for each $i = 1, 2, \ldots, n$ and any set A,
$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)} \left(= \frac{P(A|B_i)P(B_i)}{P(A)} \right) \tag{2}$$

4. **Bayes' Rule, special case:**
$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B')P(B')} \left(= \frac{P(A|B)P(B)}{P(A)} \right) \tag{3}$$
(This corresponds to the choice $B_1 = B$, $B_2 = B'$ in the general case of Bayes' Rule.)

Notes and tips

- **Memorizing Bayes' Rule:** The Total Probability Rule says that the expression appearing in the denominator in Bayes' Rule is equal to $P(A)$. If you remember this rule, you could get by memorizing the simpler version of Bayes' Rule given by the latter formulas (in parentheses) in (2) and (3). However, I recommend memorizing Bayes' Rule in the first form, since that is the form that you normally need in applications.

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

- **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".)

- **Recognizing Bayes' Rule problems:** Bayes' Rule is a formula for reversing the order in conditional probabilities. Many (but not all) conditional probability problems in the actuarial exams 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 drug test problem, the probability sought is that of someone taking drugs given that he/she tests positive, whereas the reverse conditional probability, that someone tests positive given that he/she takes drugs, is given.

- **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.

  – "5 percent *of those* taking drugs test negative."
  Translation: "P(test negative | take drugs) = 0.05."

  – "*For each smoker*, the probability of dying during the year is 0.05"
  Translation: "P(dying | smoker)=0.05"

  – "A blood test indicates the presence of a disease 95% of the time the disease is actually present."
  Translation: "P(test indicates disease | has disease)=0.95"

  – "Males *who have a circulation problem* are twice as likely to be smokers as those who do not have a circulation problem."
  Translation: "P(smoker | circulation problem)=2 P(smoker | no circulation problem)"