Lecture 18, Bayes' Theorem & Random Variables

Theorem 1 (Bayes' Theorem)**.** *Assume $\Omega = F_1 \cup F_2$, and $F_1 \cap F_2 = \emptyset$. Then*

$$p(F_i|E) = \frac{p(F_i|E)p(F_i)}{p(E|F_1)p(F_1) + p(E|F_2)p(F_2)}, \quad i = 1, 2$$

Proof. As

$$p(E \cap F_i) = p(E|F_i)p(F_i),$$

and

$$p(E) = p(E \cap F_1) + p(E \cap F_2) = p(E|F_1)p(F_1) + p(E|F_2)p(F_2).$$

Then

$$p(F_i|E) = \frac{p(E \cap F_i)}{p(E)} = \frac{p(F_i|E)p(F_i)}{p(E|F_1)p(F_1) + p(E|F_2)p(F_2)}$$

□

Example 1. *Consider a medical test for a disease D. The false positive rate is 2% and false negative rate is 1%. The probability that a people has disease D is $p(D) = 0.001$. A patient takes the test and has a positive outcome. What is the chance that he has disease D?*

We know that $p(D) = 0.001$ and $p(D^c) = 0.999$. Also we have $p(y|D) = 0.99$ and $p(y|D^c) = 0.02$, then we can compute

$$p(D|y) = \frac{p(y|D)p(D)}{p(y|D)p(D) + p(y|D^c)p(D^c)} = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.02 \times 0.999} = 0.0472$$

Trial process: repeat an experiment many times. The process is said to be independent if the outcome of the ith trial doesn't depend on the $i-1$ preceding outcomes, i.e.,

$$p(A_i|A_{i-1} \cap A_{i-2} \cap \cdots \cap A_1) = p(A_i)$$

Given an independent trail process, the probability to see an sequence of outcome $a_1, a_2, \ldots, a_n$ is given by $p(a_1)p(a_2)\cdots p(a_n)$. If the experiment has only two outcomes, the trail is also called a Bernouli trail.

Random variable: is a **function** from the sample space to the real numbers. It is not a variable, and not random.

Example 2. *Suppose that a coin is flipped three times. Let $X(t)$ be the random variable that equals the number of heads that appear when t is the outcome. Then $X(t)$ takes on the following values:*

$$X(HHH) = 3$$
$$X(HHT) = X(HTH) = X(THH) = 2$$
$$X(HTT) = X(THT) = X(TTH) = 1$$
$$X(TTT) = 0$$

People usually use notation $X = r$ to represent the set of outcomes that are mapped to value r by the random variable. Note that $X = r$ consists of outcomes, we can compute the probability of this set. Therefore, for each value of r, we can compute $p(X = r)$. The pair $\{r, p(X = r)\}$ is called the distribution of the random variable.