Lecture 2

Math 50051, Topics in Probability Theory and Stochastic Processes

We said that when trying to assign likelihoods to different events we need to look at a triple $(\Omega, eventspace, P)$. The set, collection, of all events is denoted by $\mathcal{F}$, and the triple $(\Omega, \mathcal{F}, P)$ is called the probability space.

So, an event is a subset of Ω, ie is a set formed with several ωs. To events we can assign probabilities. They are usually denoted by capital letters A, B, E, etc.

Examples:

1. Toss 4 coins. I want the event: exactly 3 heads occur.

$$A = \{(H, H, H, T), (H, H, T, H), (H, T, H, H), (T, H, H, H)\}$$

2. Toss 4 coins. I want the event: exactly 2 heads occur.

$$A = \{(H, H, T, T), (H, T, T, H), (H, T, H, T), (T, H, H, T), (T, T, H, H), (T, H, T, H)\}$$

3. Roll a die and let E be the event that we get an even number. Then $E = \{2, 4, 6\}$.
4. In the case of USDA report, depending on what is in the report, we can call it either favorable or unfavorable. Note that there might be several ωs that may lead us to call the harvest report "favorable". It is in this sense that events are collections of ωs

The collection of events we are interested in, $\mathcal{F}$, for which we will assign probabilities, might be smaller than all possible events. For example we might look at all USDA reports starting in 2000. This collection is smaller than the collection of all reports from all years. The space of events to which we assign probabilities, $\mathcal{F}$, is called in measure theoretical term a σ-field over Ω. It must obey the following conditions:

1. $\Omega \in \mathcal{F}$
2. If $A \in \mathcal{F}$ (A is an event) then $A^c \in \mathcal{F}$.
3. If $A_1, A_2, \cdots$ are events then their union must be an event, i.e. $\cup_i A_i \in \mathcal{F}$.

Examples:

1. The largest $\sigma-$field is the set of all subsets of Ω.
2. The smallest $\sigma-$field is $\mathcal{F} = \{\emptyset, \Omega\}$
3. Any Ω with only two points can have only the above as $\sigma-$fiels.

4. $\Omega = \{(H,T),(H,H),(T,H),(T,T)\}$

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{(H,T),(H,H)\}, \{(T,H),(T,T)\}\}$$

$$\mathcal{F}_2 = \{\emptyset, \Omega, \{(H,H)\}, \{(H,T),(T,H),(T,T)\}\}$$

   Observe that $B = \{(T,T)\} \notin \mathcal{F}_2$. Is it in $\mathcal{F}_1$?

5. $\mathcal{F} = \mathcal{B}(R) =$ the family of Borel sets, is the smallest $\sigma-$field containing all intervals in R.

The only item out of our triple $(\Omega, \mathcal{F}, P)$ that we have not discussed yet is P, the probability measure. The probability measure is a function on $\mathcal{F}$, the space of events, that assigns numbers to events. These numbers are the likelihoods of events occurring. They are assigned according to the following rules:

1. $P(\Omega) = 1$.
2. $0 \leq P(E) \leq 1$ for any event E.
3. for any countable collection of disjoint events we have $P(\cup_i E_i) = \sum_i P(E_i)$.

Immediate consequences:

1. If $A, B \in \mathcal{F}$ and $A \subset B$, then $P(A) \leq P(B)$. Why?
2. If $A \in \mathcal{F}$, then $P(A^c) = 1 - P(A)$. Why?
3. If $A, B \in \mathcal{F}$ then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
4. **Continuity property:** If $A_1 \supseteq A_2 \supseteq A_3 \cdots$ or if $B_1 \subseteq B_2 \subseteq B_3 \cdots$ then

$$\lim_{n \to \infty} P(A_n) = P(\cap_n A_n)$$

   in the first case, and in the second case

$$\lim_{n \to \infty} P(B_n) = P(\cup_n B_n).$$

Conditional probability

Consider the following example: The event A is: "The stock market will crash". We would like to assign a likelihood to this event. But we already know that event B occurred, where $B=$ "We are in a severe recession". The odds of a stock market crash knowing that we are in a severe recession could be represented by conditional probability.How do we formalize this on our probability space $(\Omega, \mathcal{F}, P)$?

Let $B \in \mathcal{F}$ a fixed event with positive likelihood (ie $P(B) > 0$). For any event $A \in \mathcal{F}$ we define the conditional probability measure $P_B : \mathcal{F} \to [0,1]$ by

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = P(A|B).$$

Why do we need such an object? Getting new information always allows you to update your prediction of chance. How can we relate the conditional probability P_B to the probability P?

Total probability rule: If A and B are two events such that $P(B) > 0$ then

$$P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$$

Why?

This rule extends to any partition of my sample space Ω. A partition of the sample space Ω is a sequence of events $B_1, B_2, \cdots$, with positive probability, that are disjoint and that cover my whole sample space. Then

$$P(A) = \sum_1^\infty P(A|B_i)P(B_i).$$

Bayes' rule: The other relationship between conditional expectations is called Bayes' rule. It says:

$$P(B|A) = P(A|B)\frac{P(B)}{P(A)},$$

if $P(A) > 0$. Why?

Of course, the same is true for any partition of the sample space.

Independence: 'Two events with positive probability, are independent if the occurrence of one does not influence the occurrence of the other one, ie $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

In general, two events A and B are said to be independent if

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

Why are the two definitions equivalent?

Similarly, n events are independent if

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j})$$

for $k = 2, 3, \cdots, n$ where the events $A_{i_j} \in \{A_1, \cdots, A_n\}$, for all j, are all different.