Lecture 6

Math 50051, Topics in Probability Theory and Stochastic Processes

Conditioning and σ-fields. Continuation.

Example: Suppose you are trapped in a house with 3 doors. Door 1 leads you back to the house after 1 day. Door 2 leads you back to the house after 2 days. Door 3 leads you to freedom after 3 days. On the first trial you are equally likely to pick any of the doors. If you pick door 1 or 2, then upon your return to the house, you immediately try again (until you are free).

(i) If you don't learn from your mistakes, what is the expected number of days until freedom?

(ii) Repeat part (i) with the assumption that you do learn from your mistakes.

Example Suppose $X_1, X_2, ...$ are independent identically distributed random variables with mean μ. Let S_n denote the partial sum $S_n = X_1 + ... + X_n$. Let $\mathcal{F}_n$ denote the information in $X_1, ...X_n$ $\mathcal{F}_m = \sigma(X_1, X_2, ...X_m)$, suppose $m < n$, then:

(i) $E(S_n|\mathcal{F}_m) = S_m + (n-m)\mu$

(ii) $E({S_n}^2|\mathcal{F}_m) = {S_m}^2 + (n-m)\sigma^2$

We need to move carefully in making the idea of convergence precise for stochastic processes, since random variables are functions on Ω, having distributions and moments. There are several ways in which they might be said to converge, so we begin with some simple results that will tie all these approaches together.

Markov inequality. Let X be a non-negative random variable. Then for any $a > 0$

$$P(X \geq a) \leq EX/a.$$

Chebyshov inequality. For any random variable X

$$P(|X| \geq a) \leq EX^2/a^2, \quad a > 0.$$

The event that infinitely many of the A_n occur is expressed as

$$\{A_n \; i.o.\} = \{A_n \; infinitely \; often\}$$

$$\bigcap_{n=1}^{\infty} \bigcup_{r=n}^{\infty} A_r^c.$$

Borel-Cantelli lemma. Let $(A_n; n \geq 1)$ be a collection of events, and let A be the event $\{A_n \; i.o.\}$ that infinitely many of the A_n occur. If $\sum_{n=0}^{\infty} P(A_n) < \infty$, then $P(A) = 0$.

Second Borel-Cantelli lemma. If $(A_n; n \geq 1)$ is a collection of independent events, and $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(A_n \; i.o.) = 1$.

Recall that a sequence x_n of real numbers is said to converge to a limit x, as $n \to \infty$, if $|x_n - x| \to 0$, as $n \to \infty$. Clearly, when we consider X_n with distribution $F_n(x)$, the existence of a limit X with distribution $F(x)$ must depend on the properties of the sequences $|X_n - X|$, and $|F_n(x) - F(x)|$. We therefore define the events

$$A_n(\epsilon) = \{|X_n - X| > \epsilon\}, \quad \textit{where } \epsilon > 0$$

Summation lemma. This gives a criterion for a type of convergence called *almost sure convergence*. It is straightforward to show that, as $n \to \infty$,

$$P(X_n \to X) = 1,$$

if and only if finitely many $A_n(\epsilon)$ occur, for any $\epsilon > 0$.

Convergence in probability. If, for any $\epsilon > 0$,

$$P(A_n(\epsilon)) = P(|X_n - X| > \epsilon) \to 0, \quad \textit{as } n \to \infty,$$

then X_n is said to *converge in probability* to X. We may write $X_n \xrightarrow{P} X$.

It is trivial to see that almost sure convergence implies convergence in probability; formally

$$X_n \xrightarrow{a.s.} X \Longrightarrow X_n \xrightarrow{P} X.$$

Convergence in mean square. From Chebyshov's inequality we have

$$P(|X_n - X| > \epsilon) \leq E|X_n - X|^2/\epsilon^2.$$

Therefore, if we can show that $E|X_n - X|^2 \to 0$ as $n \to \infty$, it follows that $X_n \xrightarrow{P} X$. This is often a very convenient way of showing convergence in probability, and we give it a name: if $E|X_n - X|^2 \to 0$ as $n \to \infty$, then X_n is said to *converge in mean square* to X. We may write $X_n \xrightarrow{m.s.} X$.

However, even this weaker form of convergence sometimes fails to hold; in the last resort we may have to be satisfied with showing convergence of the distributions $F_n(x)$. This is a very weak form of convergence, as it does not even require the random variables X_n to be defined on a common probability space.

Convergence in distribution. If $F_n(x) \to F(x)$ at all the points x such that $F(x)$ is continuous, then X_n is said to *converge in distribution*. We may write $X_n \xrightarrow{D} X$.

Central limit theorem. If $EX_r = \mu$ and $0 < var X_r = \sigma^2 < \infty$, then, as $n \to \infty$,

$$P\big(\frac{S_n - n\mu}{(n\sigma^2)^{1/2}} \leq x\big) \to \Phi(x),$$

where $\Phi(x)$ is the standard normal distribution.

Weak law of large numbers. If $EX_r = \mu < \infty$, then for $\epsilon > 0$, as $n \to \infty$,

$$P\Big(\Big|\frac{S_n}{n} - \mu\Big| > \epsilon\Big) \to 0.$$

Strong law of large number. As $n \to \infty$,

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu$$

for some finite constant μ, if and only if $E|X_r| < \infty$, and then $\mu = EX_1$.

The central limit theorem is the principal reason for the appearance of the normal (or 'bell-shaped') distribution in so many statistical and scientific contexts. The first version of this theorem was proved by Abraham de Moivre before 1733.