Reading Note 10A
Oct 27, 2015

Law of Large Numbers

The Law of Large Number is covered in chapter 7.1 and 7.2 in the textbook. There are two important concepts for WLLN. One is WHY, two is HOW. The WHY question is answered in textbook 7.2 the first few paragraphs. You may also consult Bertsekas and Tsitsiklis chapter 5.2 for the discussion. Now, the HOW question is done by the Markov and Chebyshev inequality., which is covered in textbook chapter 4.6. You can also read Bertsekas and Tsitsiklis chapter 5.1 for alternative discussions (which is better in my opinion.) Once you have acquired Markov and Chebyshev inequality, the proof of WLLN becomes immediate. Textbook pp. 366 has the full proof. I would also encourage you to read Bertsekas and Tsitsiklis Example 5.4 and Example 5.5 to gain a better sense of how WLLN can help in practice.

A major challenge in WLLN is the concept of convergence in probability. This is a difficult concept if you have not seen it before. Basically, in classical analysis the convergence is well defined for a sequence of real valued numbers \( y_1, y_2, y_3, \ldots \). However, if we consider a sequence of random variables \( Y_1, Y_2, Y_3, \ldots \), the convergence becomes less clear because a random variable is NOT a real value. A random variable is determined by its PDF. All we see are samples taken from this PDF. So in order to describe the convergence of random variables, we need to define it using some probability measure. The convergence mentioned in the WLLN is one type. It is called convergence in probability. If you read Bertsekas and Tsitsiklis chapter 5.3 they will have a formal definition.

Now, how about Strong Law of Large Number? SLLN is a much more powerful statement than WLLN. The difference of the two is however very subtle:

\[
\text{WLLN: } \lim_{n \to \infty} \mathbb{P}[|M_n - \mu| > \varepsilon] = 0 \quad \text{SLLN: } \mathbb{P} \left[ \lim_{n \to \infty} |M_n - \mu| > \varepsilon \right] = 0.
\]

In a nutshell, what matters is the position of the limit and the probability. But this small change has drastic different meaning. In WLLN, \( \mathbb{P}[|M_n - \mu| > \varepsilon] \) is the probability of certain event and so it is a real number (between 0 and 1). So the limit of this sequence of real numbers is well defined by the classical convergence theory. In SLLN, \( \lim_{n \to \infty} |M_n - \mu| \) requires us to determine a limiting object of a sequence of random variables. This limiting object itself is a random variable (otherwise we cannot take probability outside). Finding this limiting object is difficult, which makes SLLN difficult to prove.

The type of convergence in SLLN is called almost sure convergence. Proving almost sure convergence is often more challenging than proving probabilistic convergence. You are welcome to read Bertsekas and Tsitsiklis chapter 5.5 for more discussion. In my opinion this chapter 5.5 is the best discussion among all textbooks I have seen. Very precise and very clear.