Statistics 4657/9657 : Strong Law of Large Numbers

The text gives an interesting proof of the Strong Law of Large Numbers (SLLN). Here the classic proof is given. This is done since it involves a number of tools that are useful in other settings, such as equivalent sequences and truncation. This proof is not as elegant as the one given in the text.

Reference : R Durrett, Probability : Theory and Examples. See Section 1.8, p 81, and parts of Section 1.7.

Theorem 1 (Kolmogorov's Inequality) *Suppose $X_1, \ldots, X_n$ are independent r.v.s with $E(X_j) = 0$ and finite $Var(X_j) = \sigma_j^2$. Let $S_k = X_1 + \ldots + X_k$. Then*

$$P\left(\max_{1 \le k \le n} |S_k| \ge x\right) \le \frac{Var(S_n)}{x^2}$$

Theorem 2 *Suppose X_j are independent r.v.s with $E(X_j) = 0$. If $\sum_{j=1}^{\infty} Var(X_j) < \infty$ then $\sum_{j=1}^{\infty} X_j$ exists almost surely.*

Remark : This is proven by showing that $S_n = \sum_{j=1}^{n} X_j$ is almost surely a Cauchy sequence.

Proof Let $W_M = \max_{n,m \ge M} |S_m - S_n|$.

For $M < N$

$$P(\max_{M \le N} |S_m - S_M| \ge \epsilon) \quad \le \quad \frac{\mathrm{Var}(S_N - S_M)}{\epsilon^2}$$

Letting $N \to \infty$ then

$$\begin{aligned} P(\max_{m \ge M} |S_m - S_M| \ge \epsilon) &= \lim_{N \to \infty} P(\max_{M \le N} |S_m - S_M| \ge \epsilon) \\ &\le \lim_{N \to \infty} \frac{\sum_{j=M}^{\infty} \sigma_j^2}{\epsilon^2} < \infty \ . \end{aligned}$$

Notice that if $\max_{m \ge M} |S_m - S_M| \ge \epsilon$ then for all $n, m \ge M$ we have

$$|S_n - S_m| \le |S_n - S_M| + |S_m - S_M| \le 2\epsilon$$

Thus $P(W_M \geq 2\epsilon) \leq P(\max_{m \geq M} |S_m - S_M| \geq \epsilon)$.

Let $C = \{\omega : S_n(\omega) \text{ converges }\}$. If $\omega \notin C$, then there exists $\epsilon > 0$ and for any $M \geq 1$, there exists $m \geq M$, such that $|S_m(\omega) - S_\infty(\omega)| \geq \epsilon$. Thus

$$C^c \subseteq \cap_{k \geq 1} \cup_{m=M}^{\infty} A_{M,k} = \cap_{k \geq 1} \cup_{m=M}^{\infty} \{\omega : W_M(\omega) \geq 1/k\}$$

Thus for fixed $\epsilon = 1/k$, and for any $M \geq 1$

$$P(C^c) \leq P(A_{M,k}) \leq \frac{1}{1/k} \sum_{m=M}^{\infty} \sigma_j^2 \ .$$

Thus $P(C^c) \leq k \limsup_M P(A_{M,k}) = 0$, and hence $P(C) = 1$. Thus S_n converges almost surely.

END of PROOF of Theorem 2.

Theorem 3 (Kronecker's Lemma) *Suppose $a_n \geq 0$ and a_n monotonically increases to ∞. Suppose also that $\sum_{j=1}^{\infty} x_j/a_j$ converges. Then*

$$\frac{1}{a_n} \sum_{j=1}^{m} x_j \to 0 \ .$$

Proof : Let $a_0 = b_0 = 0$ and for $n \geq 1$

$$b_n = \sum_{j=1}^{n} \frac{x_j}{a_j}$$

Then $x_n = a_n(b_n - b_{n-1})$ for $n \geq 1$. By the hypothesis $b_n \to b_\infty$ for some finite limit. Also

$$\frac{1}{a_n} \sum_{j=1}^{n} x_n = \frac{1}{a_n} \left\{ \sum_{j=1}^{n} a_j b_j - \sum_{j=1}^{n} a_j b_{j-1} \right\} = b_n - \sum_{j=1}^{n} \frac{(a_j - a_{j-1})}{a_n} b_j \to b_\infty - b_\infty$$

Aside For the sum in the second last part, the student should use an idea similar to a Cesaro mean to show it converges to b_∞.

Theorem 4 *Suppose Y_j are independent random variables with $E(Y_j) = 0$. Let $T_n = Y_1 + \ldots + Y_n$. If $a_n > 0$ increases to ∞ and $\sum_{j=1}^{\infty} E(Y_j^2)/a_j^2 < \infty$ then*

$$\frac{T_n}{a_n} \to 0 \ a.s. \ as \ n \to \infty \ .$$

Proof :

By Theorem 2, there exists an A such that $P(A) = 1$ and $\forall \omega \in A$, $\sum_{j=1}^{n} \frac{Y_j(\omega)}{a_j}$ converges.

Thus by Theorem 3, for $\omega \in A$,

$$\frac{1}{a_n} \sum_{j=1}^{n} Y_j(\omega) \to 0$$

Theorem 5 (SSLN) *Suppose $X_j, j \geq 1$ are iid r.v.s with $E(X_1) = \mu$ finite. Write $S_n = X_1 + \ldots + X_n$. Then $\frac{S_n}{n} \to \mu$ a.s. as $n \to \infty$.*

Proof.

Define $Y_j = X_j \mathrm{I}(|X_j| \leq j)$.

Notice that $P(X_j \neq Y_j) = P(|X_1| > j)$ and hence

$$\sum_{j \geq 1} P(X_j \neq Y_j) = \sum_{j \geq 1} P(|X_1| > j) \leq \mathrm{E}(|X_1|) < \infty.$$

Thus $\{X_j\}$ and $\{Y_j\}$ are equivalent sequences.

Next show that

$$\sum_{j \geq 1} \frac{\mathrm{Var}(Y_j)}{j^2} < \infty \ .$$

Let $\mu_j = \mathrm{E}(Y_j)$

$$\sum_{j \geq 1} \frac{\mathrm{Var}(Y_j)}{j^2} \leq \sum_{j \geq 1} \frac{\mathrm{E}(Y_j^2)}{j_2}$$

Note that $Y_j^2 = X_j^2 I(|X_j| \leq j)$, thus for any $y \geq 0$, if $X_j^2 < y$ then $Y_j^2 < y$. Thus $P(X_j^2 < y) \leq P(Y_j^2 < y)$ and hence $P(|X_1| < y) = P(|X_j| < y) \leq P(|Y_j| < y)$ $P(|X_1| \geq y) \leq P(|Y_j| \geq y)$. Thus

$$\mathrm{E}(Y_j^2) = \int_0^\infty 2y dP(|Y_j| \geq y) \leq \int_0^j 2y P(|X_1| \geq y) \ .$$

Therefore

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{\mathrm{E}(Y_j^2)}{j^2} &\leq \sum_{j=1}^{\infty} j^{-2} \int_0^j 2y P(|X_1| \geq y) \\ &= \int_0^\infty \sum_{j: j \geq y, j \geq 1} \frac{1}{j^2} 2y P(|X_1| \geq y) \end{aligned}$$

$$
\begin{aligned}
&\le \int_0^\infty \frac{2}{y} 2y P(|X_1| \ge y) \\
&\le 4\int_0^\infty P(|X_1| \ge y) \\
&= 4E(|X_1|) < \infty \ .
\end{aligned}
$$

Thus

$$\frac{1}{n}\sum_{j=1}^{n}(Y_j - \mu_j) \to 0 \text{ a.s.}$$

Notice that $\mu_j \to \mathrm{E}(X_1) = \mu$ as $j \to \infty$, thus by the Cesaro means property $\frac{1}{n}\sum_{j=1}^{n}\mu_j \to \mu$.

Thus

$$
\begin{aligned}
\frac{1}{n}\sum_{j=1}^{n} X_j &= \sum_{j=1}^{n}(X_j - Y_j) + \frac{1}{n}\sum_{j=1}^{n}(Y_j - \mu) + \frac{1}{n}\sum_{j=1}^{n}\mu_j \\
&\to 0 + 0 + \mu \text{ a.s.}
\end{aligned}
$$