Math 617: Day 49

We need one more lemma before proving the Strong Law of Large Numbers:

Kolmogorov's Inequality: Let $X_1, \dots, X_n$ be independent, real-valued r.v.'s w/ mean 0 & variances $\sigma_1^2, \dots \sigma_n^2$.

For each $k=1,\dots,n$, let $S_k = X_1 + \dots + X_k$. For $\varepsilon > 0$,

$$P\left(\max_k |S_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^n \sigma_i^2 .$$

Pf: Notice that $\max_k |S_k| \geq \varepsilon$ is really the union of a bunch of disjoint events: for some k, S_k will be the first partial sum to escape from $(-\varepsilon, \varepsilon)$. If we let $A_k = \{ |S_k| \geq \varepsilon \ \& \ |S_j| < \varepsilon \ \forall j < k\}$, then

$$P\left(\max_k |S_k| \geq \varepsilon\right) = P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) \quad \text{b/c the } A_k \text{ are pairwise disjoint.}$$

Now, we're trying to show this is $\leq \frac{1}{\varepsilon^2} \sum \sigma_i^2 = \frac{1}{\varepsilon^2} \sigma^2(S_n) = \frac{1}{\varepsilon^2} E(S_n^2) - E|S_n|^2 = \frac{1}{\varepsilon^2} E(S_n^2)$ since the X_i have mean 0.

But then $S_n^2 \geq S_n^2 \sum_{k=1}^n \chi_{A_k}$ since $0 \leq \sum \chi_{A_k} \leq 1$, so

$$E(S_n^2) \geq E\left(\sum S_n^2 \chi_{A_k}\right) = \sum E(S_n^2 \chi_{A_k}) .$$

Now we use the fact that $S_n^2 = (S_k + (S_n - S_k))^2 = S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2$ to re-write the above as

$$E(S_n^2) \geq \sum E\left((S_k^2 + 2S_k(S_n-S_k) + (S_n-S_k)^2)\chi_{A_k}\right) = \sum E(S_k^2 \chi_{A_k}) + \underbrace{2\sum E(S_k(S_n-S_k)\chi_{A_k})}_{0} + \underbrace{\sum E((S_n-S_k)^2 \chi_{A_k})}_{\geq 0}$$

$$\geq \sum \varepsilon^2 P(A_k) = \varepsilon^2 \sum P(A_k)$$

since $S_k^2 \chi_{A_k} \geq \varepsilon^2 \chi_{A_k} \ \forall k$. So we're done. ∎

Kolmogorov's Strong Law of Large Numbers: Let $(X_n)_{n=1}^\infty$ be a seq. of independent, real-valued, L^2 r.v.'s w/ means (μ_n) & variances (σ_n^2) w/ $\sum_{n=1}^\infty \frac{\sigma_n^2}{n^2} < \infty$. Then

$$\frac{1}{n} \sum_{i=1}^n (X_i - \mu_i) \longrightarrow 0 \quad \text{almost surely (a.k.a. almost everywhere) as } n \to \infty.$$

Proof: Let $S_n = \sum_{i=1}^n X_i - \mu_i$. Then we're trying to show that, for any $\varepsilon > 0$, $P\left(\limsup \frac{1}{n}|S_n| < \varepsilon\right) = 1$.

But this will follow if we know that $\frac{1}{n}|S_n| \geq \varepsilon$ for only finitely many n.

The idea is to use the Borel-Cantelli lemma, which says that, if A_k are events w/ $\sum P(A_k) < \infty$, then only finitely many of the events occur simultaneously.

So we want to let A_k be something like the event that $\frac{1}{k}|S_k| \geq \varepsilon$. Of course, if we just let A_k be that, we won't be able to guarantee that $\sum P(A_k) < \infty$.

Instead, we let A_k be the event that $\frac{1}{n}|S_n| \geq \varepsilon$ for some $2^{k-1} \leq n < 2^k$. In other words, on A_k we have

$|S_n| \geq n\varepsilon \geq 2^{k-1}\varepsilon$ for some $n < 2^k$, which means $\max_{1 \leq n < 2^k} |S_n| \geq 2^{k-1}\varepsilon$.

Then Kolmogorov's Inequality implies

$$P(A_k) \leq P\Big(\max_{1 \leq n < 2^k} |S_n| \geq 2^{k-1}\varepsilon\Big) \leq \frac{1}{(2^{k-1}\varepsilon)^2} \sum_{n=1}^{2^k} \sigma_n^2 .$$

Now we can sum & expand to get a convergent series:

$$\sum_{k=1}^{\infty} P(A_k) \leq \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} \sum_{n=1}^{2^k} \frac{1}{2^{2k-2}} \sigma_n^2 = \frac{4}{\varepsilon} \sum_{k=1}^{\infty} \sum_{n=1}^{2^k} \frac{1}{2^{2k}} \sigma_n^2 = \frac{4}{\varepsilon}\Big[\frac{1}{2^2}(\sigma_1^2+\sigma_2^2) + \frac{1}{2^4}(\sigma_1^2+\sigma_2^2+\sigma_3^2+\sigma_4^2) + \dots\Big]$$

$$= \frac{4}{\varepsilon}\Big[\Big(\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\Big)\sigma_1^2 + \Big(\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\Big)\sigma_2^2 + \Big(\sum_{k=2}^{\infty}\frac{1}{2^{2k}}\Big)\sigma_3^2 + \Big(\sum_{k=2}^{\infty}\frac{1}{2^{2k}}\Big)\sigma_4^2 + \dots\Big]$$

$$= \frac{4}{\varepsilon} \sum_{n=1}^{\infty} \Big(\sum_{k \geq \log_2 n} \frac{1}{2^{2k}}\Big) \sigma_n^2$$

$$= \frac{4}{\varepsilon} \sum_{n=1}^{\infty} \Big(\frac{1}{3} \cdot \frac{1}{2^{2\lfloor \log_2 n \rfloor - 2}}\Big) \sigma_n^2$$

$$\leq \frac{4}{\varepsilon} \sum_{n=1}^{\infty} \Big(\frac{1}{3} \frac{4}{2^{2\log_2 n}}\Big) \sigma_n^2$$

$$= \frac{16}{3\varepsilon} \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2}$$

which is finite by hypothesis, so we're done after applying Borel-Cantelli. ∎

We get to make weaker assumptions on the regularity of the X_i if they're identically distributed:

Khinchine's Strong Law of Large Numbers: If $(X_n)_{n=1}^{\infty}$ is a seq. of i.i.d. real-valued L^1 r.v.'s w/ mean μ, then

$\frac{1}{n} \sum_{i=1}^{n} X_i \to \mu$ almost surely as $n \to \infty$.