EECS 126 — 2.16 LECTURE 9

From last time: Chebyshev's Inequality implies WLLN. Also, WLLN justifies probability in that axioms are compatible with frequency and empirical frameworks. Our proof for WLLN operated on expectation, without computing distributions/probabilities.

CHERNOFF BOUNDS

For a RV X and an $a \in \mathbb{R}$, $\Pr\{X \geq a\} \leq \frac{E[e^{tX}]}{e^{ta}} = e^{-ta} M_X(t), \quad t > 0$
Since this holds for all $t > 0$, we can optimize.

PROOF $\Pr\{X \geq a\} = \Pr\{tX \geq ta\} = P\{e^{tX} \geq e^{ta}\} \leq \frac{E[e^{tX}]}{e^{ta}}$ ▨

IDEA Chebyshev gives "better" bounds than Markov because it uses $Var(X) < \infty$. Chernoff, using MGF, uses info about <u>all</u> moments.

EXAMPLE $Z \sim N(0,1)$. We know that Φ has no closed form, but Chernoff gives us good control of tail probabilities.

$$\Pr\{Z \geq a\} \leq e^{-ta} M_Z(t) = e^{-ta} e^{t^2/2} = e^{t^2/2 - ta} \longrightarrow d/dt\, e^{t^2/2 - ta} = (t-a)e^{t^2/2 - ta} = 0$$

To optimize/minimize, set $t = a \longrightarrow \Pr\{Z \geq a\} \leq e^{-a^2/2} \quad a > 0$

CONVERGENCE

Language of "limits" in probability. Given a sequence of RVs $X_1, X_2, \ldots$, what does it mean to say $\lim_{n\to\infty} X_n = X$? Nothing, without more information.

RVs are functions, so we need to specify type of convergence.
Pointwise convergence : $\lim_{n\to\infty} f_n(x) = f(x) \quad \forall x$
In L_1-norm : $\lim_{n\to\infty} \int |f_n(x) - f(x)|\, dx = 0$

<u>Three modes:</u>

1. Almost Sure. $X_n \to X$ almost surely if $P(\lim_{n\to\infty} X_n = X) = 1$.
2. In Probability. $X_n \to X$ in probability if $\forall \varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$.
3. In Distribution. $X_n \to X$ in distribution if $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ $\forall$ continuity points x of F_X.

$(X_n \to X \text{ a.s.}) \Longrightarrow (X_n \to X \text{ in prob.}) \Longrightarrow (X_n \to X \text{ in dist.})$
All of these implications are strict.

Proofs are a bit technical, feel free to Google.

Strong Law of Large Numbers

If $X_1, X_2, \ldots \sim_{IID} X$ and $E[X] < \infty$, then $\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow E[X]$ a.s.

Same claim as WLLN, but almost surely instead of in probability.
Also tells us that individual sample paths $\frac{1}{n}\sum X_i(\omega) = E[X]$ $\forall \omega$ if $P=1$.
Doesn't tell us anomalies can't happen, but that they won't happen.

Central Limit Theorem

Let $X_1, X_2, \ldots \sim_{IID} X$, and $Var(X) = \sigma^2 < \infty$, $E[X] = \mu$.
Define $S_n = \left(\sum_{i=1}^{n} X_i - \mu\right) / \sigma\sqrt{n}$. Then, $S_n \longrightarrow Z \sim N(0,1)$ in distribution. Also,
$\lim_{n\to\infty} \Pr\{S_n \le x\} = \Phi(x)$ $\forall x \in \mathbb{R}$.

Proof is in lecture slides.

Caution: $X_n \longrightarrow X$ does not imply $E[X_n] \to E[X]$ without more information.