Chapter 5

Limit Theorems

5.1 Markov and Chebychev Inequalities

Markov Inequality is a -typically loose- bound on the value of a nonnegative random variable X with a known mean $E(X)$.

Markov Inequality: If a random variable X can take only nonnegative values, then

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Proof:

Ex: Let X be uniform in $[5, 10]$. Compute probabilities that X exceeds certain values and compare them with the bound given by Markov Inequality.

To be able to bound probabilities for general random variables (not necessarily positive), and to get a tighter bound, we can apply Markov Inequality to $(X - E(X))^2$ and obtain:

Chebychev Inequality: For random variable X with mean $E(X)$ and variance σ^2, and any real number $a > 0$,

$$P(|X - E(X)| \geq a) \leq \frac{\sigma^2}{a^2}$$

Proof: Bound $P((X - E(X))^2 \geq a^2$ using Markov Inequality.

Note that Chebychev's Inequality uses more information about X in order to provide a tighter bound about the probabilities related to X. In addition to the mean (a first-order statistic), it also uses the variance, which is a second-order statistic. You can easily imagine two very different random variables with the same mean: for example, a zero-mean Gaussian with variance 2, and a discrete random variable that takes on the values +0.1, and -0.1 equally probably. Markov Inequality does not distinguish between these distributions, where as Chebychev Inequality does.

It is possible to get better bounds that use more information- such as the Chernoff bound (but those are beyond the scope of this course.)

Ex: Use Chebychev's Inequality to lower-bound the probability that a Gaussian within two standard deviations of its mean.

5.2 Probabilistic Convergence

Let us remember the definition of convergence for a deterministic sequence of numbers from basic Calculus. Let $\{a_n\}$ be a sequence of numbers, indexed by n. $\lim_{n\to\infty} a_n = a$ means, for every $\epsilon > 0$, there exists an n_o such that for all $n > n_o$, $|a_n - a| < \epsilon$.

5.2.1 Convergence "In Probability"

Now, instead of a sequence of numbers, consider a sequence of random variables $\{Y_n\}$, indexed by n. Y_n converges in probability to a number a if for every $\epsilon > 0$, $\lim_{n\to\infty} P(|Y_n - a| > \epsilon) = 0$. The notation is: $Y_n \stackrel{i.p}{\to} a$.

Ex: Suppose for each n, Y_n takes the value n with probability $1/n$, and the value zero with probability $1 - 1/n$. Does $\{Y_n\}$ converge, and if so, to what?