MATH 589

The Central Limit Problem—an overview

The first Central Limit Theorem (in this case, the weak law of large numbers) was proven by Bernoulli in 1713. Let $0 < p < 1$, $q = 1 - p$, and let X_n be a sequence of i.i.d. random variables such that

$$P(X_n = 1) = p, \qquad P(X_n = 0) = q.$$

Let $S_n = X_1 + \cdots + X_n$. The random variables S_n have the binomial distribution

$$P(S_n = j) = \binom{n}{j} p^j q^{n-j}, \qquad j = 0, \cdots, n.$$

Bernoulli has proven that

$$\frac{S_n - np}{n} \to 0 \qquad \text{in probability.}$$

The Classical degenerate convergence condition discussed in the previous lecture notes gives the optimal generalization of the Bernoulli result to an arbitary sequence of independent random variables.

In 1732 de Moivre has proven that

$$P_n(x) = P(S_n = j) \sim \frac{1}{\sqrt{2\pi npq}} \mathrm{e}^{-x^2/2}, \qquad x = \frac{j - np}{\sqrt{npq}},$$

uniformly on every finite interval $[a, b]$ of values of x. In 1801, Laplace has extended the result and showed that

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right) = (2\pi)^{-1/2} \int_a^b \mathrm{e}^{-x^2/2} \mathrm{d}x.$$

Exercise: Prove the de Moivre-Laplace theorem using only Stirling's formula!

Obviously, the Lindberg-Feller theorem discussed in the previous lecture notes gives an (essentially) optimal generalization of the de Moivre-Laplace theorem.

In 1802, Laplace modifed the Bernoulli case by assuming that $p = p_n$ depends upon a total number n of the trials in such a manner that $np_n \to \lambda > 0$. More precisely, Laplace considered a double sequence of random variables X_{nk}, $1 \le k \le n$, such that for fixed n X_{nk} are i.i.d. and

$$P(X_{nk} = 1) = p_n, \qquad P(X_{nk} = 0) = q_n,$$

where $\lim_{n\to\infty} np_n = \lambda > 0$. Let $S_n = \sum_{k=1}^{n} X_{nk}$. A simple analysis (done in class) yields that

$$\lim_{n\to\infty} P(S_n = j) = \mathrm{e}^{-\lambda}\frac{\lambda^j}{j!}, \qquad j = 0, 1, \cdots$$

While Bernoulli and de Moivre-Poisson results played a central role in the development of the probability theory, the Poisson law for many years stood isolated and ignored. In fact, so far we have only briefly discussed the Poisson law in the course. It will turn out, however, that the Poisson law is more fundamental to the Central Limit Problem/Theorem then the two others.

The proper formulation of the Central Limit Problem involves a double sequence X_{nk}, $1 \le k \le k_n$, $k_n \to \infty$, of row-wise independent random variables. Set

$$S_n = \sum_{k=1}^{k_n} X_{nk}.$$

The Central Limit problems consists of two parts:

Part I. Find the family of all possible limit laws of these sums.

Part II. Find conditions for convergence to any specified law in this family.

Of course, the Bernoulli case (and any weak law of large numbers), the de Moivre-Laplace theorem (and the Lindberg-Feller theorem), and also the Poisson law can be naturally formulated in the above framework. For example, in the Central degenenarate convergence condition one takes $k_n = n$, $X_{nk} = X_k/n$. In the Lindberg-Feller theorem, one takes $k_n = n$ and $X_{nk} = (X_k - \mathbb{E}(X_k))/s_n$.

The Central Limit Problem, as we have formulated it, is obviously too general. The natural restriction is to consider the summands X_{nk} that are uniformly small. More precisely, the double sequence is called infinitesimal if for every $\epsilon > 0$,

$$\lim_{n\to\infty} \max_{1\le k\le k_n} P(|X_{nk}| > \epsilon) = 0.$$

Under this condition the Central Limit Problem admits a very elegant solution in which the fundamental role is played by the infinitely divisible random variables.

In what follows I will describe the solution to the Central Limit Problem under the so called bounded variance condition. In this case, I will give complete proofs. I will also describe (without proofs) the general solution to the Central Limit Problem. The proofs in the general case follow the same ideas as in the bounded variance case. However, they are rather lengthy and (given how late we are with the material) I will skip them. A fairly complete discussion of the solution to the general Central Limit Problem can be found in Burrill's book.