Math 464/564, O. Knill, **April 1997**

The central limit theorem

DEFINITIONS. A sequence of random variables X_n **converges in distribution** to X if $\mathrm{P}[X_n \leq t] \to \mathrm{P}[X \leq t]$ for all $t \in \mathbb{R}$. Denote by $Y^* = (Y - \mathrm{E}[Y])/\sigma(Y)$ the **normalized random variable** to Y.

THE CENTRAL LIMIT THEOREM FOR INDEPENDENT RANDOM VARIABLES. Assume X_n are independent random variables satisfying $\sup_n \mathrm{E}[|X_n|^3] < \infty, \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathrm{Var}[X_i] > 0$. Then S_n^* converges in distribution to a standard normal distributed random variable. The same conclusion holds if X_n are IID with nonzero finite variance.

PROOF. Define for fixed $n \geq 1$ the random variables

$$Y_i = (X_i - E[X_i])/\sigma(S_n),\ 1 \leq i \leq n$$

so that $S_n^* = \sum_{i=1}^n Y_i$. Define $N(0, \sigma^2 = \mathrm{Var}[Y_i])$-distributed random variables $\tilde{Y}_i$ having the property that $\{Y_1, \dots, Y_n, \tilde{Y}_1, \dots \tilde{Y}_n\}$ are independent. The distribution of $\tilde{S}_n = \sum_{i=1}^n \tilde{Y}_i$ is just the normal distribution $N(0,1)$. The convergence in distribution is equivalent to $E[f(S_n^*)] - E[f(\tilde{S}_n)] \to 0$ for any smooth function f which is zero outside some interval. Define

$$Z_k = \tilde{Y}_1 + \dots \tilde{Y}_{k-1} + Y_{k+1} + \dots + Y_n\ .$$

Note that $Z_1 + Y_1 = S_n^*$ and $Z_n + \tilde{Y}_n = \tilde{S}_n$. Using a telescopic sum and Taylor's theorem, we write

$$\begin{aligned} f(S_n^*) - f(\tilde{S}_n) &= \sum_{k=1}^n [f(Z_k + Y_k) - f(Z_k + \tilde{Y}_k)] \\ &= \sum_{k=1}^n [f'(Z_k)(Y_k - \tilde{Y}_k) + \frac{1}{2} f''(Z_k)(Y_k^2 - \tilde{Y}_k^2) \\ &\quad + R(Z_k, Y_k) + R(Z_k, \tilde{Y}_k)] \end{aligned}$$

with a Taylor rest term $R(Z,Y)$ depending on f. We get therefore

$$|E[f(S_n^*)] - E[f(\tilde{S}_n)]| \leq \sum_{k=1}^n E[|R(Z_k, Y_k)|] + E[|R(Z_k, \tilde{Y}_k)|]\ . \tag{1}$$

Since $\tilde{Y}_k$ are $N(0,\sigma^2)$-distributed we get using the computation of the moments of normal random variables and Jensen's inequality

$$E[|\tilde{Y}_k|^3] = \sqrt{\frac{8}{\pi}}\sigma^3 = \sqrt{\frac{8}{\pi}} E[|Y_k|^2]^{3/2} \leq \sqrt{\frac{8}{\pi}} E[|Y_k|^3]\ .$$

The Taylor rest can be estimated as $|R(Z_k, Y_k)| \leq \mathrm{const} \cdot |Y_k|^3$ so that

$$\begin{aligned} \sum_{k=1}^n E[|R(Z_k, Y_k)|] + E[|R(Z_k, \tilde{Y}_k)|] &\leq \mathrm{const} \cdot \sum_{k=1}^n E[|Y_k|^3] \\ &\leq \mathrm{const} \cdot n \cdot \sup_i \mathrm{E}[|X_i|^3]/\mathrm{Var}[S_n]^{3/2} \\ &= \mathrm{const} \cdot \frac{\sup_i \mathrm{E}[|X_i|]^3}{(\mathrm{Var}[S_n]/n)^{3/2}} \cdot \frac{1}{\sqrt{n}} \to 0\ . \end{aligned}$$

If we know only the finite nonzero variance of X_n but instead that the X_n are IID, we estimate the Taylor rest as $|R(z,y)| \leq \delta(y) \cdot y^2$, where $\delta(y) \to 0$ for $|y| \to 0$. The IID property and the dominated convergence gives

$$\begin{aligned} \sum_{k=1}^n E[|R(Z_k, Y_k)|] + E[|R(Z_k, \tilde{Y}_k)|] &\leq \sum_{k=1}^n E[\delta(Y_k) Y_k^2] + E[\delta(\tilde{Y}_k)\tilde{Y}_k^2] \\ &= n \cdot E[\delta(\frac{X_1}{\sigma\sqrt{n}})\frac{X_1}{\sigma\sqrt{n}}] + n \cdot E[\delta(\frac{\tilde{X}_1^2}{\sigma^2 n})\frac{\tilde{X}_1^2}{\sigma^2 n}] \to 0\ . \end{aligned}$$