Central Limit Theorem

Will Perkins

February 14, 2013

Approximating Binomial Probabilities

What is

$$\Pr[Bin(n, 1/2) = n/2 + x\sqrt{n/2}]$$

Exact:

$$= \binom{n}{n/2 + x\sqrt{n/2}} 2^{-n}$$

Then use Stirling's formula.

Reproving the WLLN

Theorem

Let $X_1, X_2, \ldots$ be iid with $\mathbb{E}X_i = \mu$. Then

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{D} \mu$$

Proof: Let $U_n = \frac{X_1+X_2+\cdots+X_n}{n}$.
Calculate $\phi_{U_n}(t)$ and show it converges to $e^{it\mu}$ as $n \to \infty$.

$$\phi_{U_n}(t) = \phi_{X_1}(t/n)^n$$

Reproving the WLLN

We know that $\phi_{X_1}(0) = 1$ and that it is a uniformly continuous function, so this suggests using a Taylor expansion, since t/n is very close to 0 for large n.

$$\phi_{X_1}(t/n) = \phi_{X_1}(0) + \frac{t}{n}\phi'_{X_1}(0) + o\left(t/n\right)$$

$$= 1 + it\mu/n + o(t/n)$$

Now we raise this to the nth power, and use a familiar limit:

$$\phi_{U_n}(t) = \phi_{X_1}(t/n)^n = (1 + it\mu/n + o(t/n))^n \to e^{it\mu}$$

as $n \to \infty$.
$e^{it\mu}$ is the characteristic function of the constant μ, so we've shown that

$$U_n \xrightarrow{D} \mu$$

The Central Limit Theorem

Theorem

Let $X_1, X_2, \dots$ be iid random variables with mean μ and variance σ^2. Then

$$\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{\sigma^2 n}} \xrightarrow{D} N(0, 1)$$

as $n \to \infty$.

Proof

We'll use a similar proof to the proof of the WLLN.
Let

$$U_n = \frac{X_1 + \cdots + X_n - n\mu}{\sqrt{\sigma^2 n}}$$

Calculate:

$$\phi_{U_n}(t) = \phi_{X_1 - \mu}(t/\sigma\sqrt{n})^n$$

Proof

Taylor expansion around 0:

$$\phi_{X_1-\mu}(t/\sigma\sqrt{n}) = \phi_{X_1-\mu}(0) + \frac{t}{\sigma\sqrt{n}}\phi'_{X_1-\mu}(0) + \frac{t^2}{2\sigma^2 n}\phi''_{X_1-\mu}(0) + o\left(t^2/n\right)$$

$$= 1 + 0 - \frac{t^2}{2n} + o(t^2/n)$$

And so, raising this to the *n*th power gives:

$$\phi_{U_n}(t) = (1 - \frac{t^2}{2n} + o(t^2/n))^n \to e^{-t^2/2}$$

as $n \to \infty$, and $e^{-t^2/2}$ is the characteristic function of a $N(0,1)$ RV.

History of the CLT

Abraham deMoivre, 1718, *The Doctrine of Chance*.
At this point there was no Gaussian distribution (and no Gauss), and no Fourier transform (no Fourier).

DeMoivre spent 3 years in prison in France for religious reasons, then moved to England at age 21 where he met Isaac Newton. He was interested in calculating the odds in gambling games, and in particular wanted an approximation to Binomial probabilities.

Lindeberg-Feller CLT

For non-identically distributed rv's. X_i's independent with mean 0 and finite variances.
Let $s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$. Let F_n be the distribution function of X_n.
The Lindeberg condition is that for all $\epsilon > 0$:

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^{n} \int_{|x|>\epsilon s_j} x^2 \, dF_j(x) = 0$$

Then

$$\frac{X_1 + \cdots + X_n}{s_n} \xrightarrow{D} N(0,1)$$

CLT for Triangular Arrays

What's a triangular array?

$$\begin{pmatrix} X_{1,1} & & & \\ X_{2,1} & X_{2,2} & & \\ X_{3,1} & X_{3,2} & X_{3,3} & \\ X_{4,1} & X_{4,2} & X_{4,3} & X_{4,4} \\ \dots & & & \dots \end{pmatrix}$$

Let $S_n = \sum_{j=1}^{n} S_{n,j}$ (Sum of row n).
What are conditions on the random variables so that a properly centered and normalized S_n converges to a normal distribution as $n \to \infty$?

CLT for Triangular Arrays

1. Independence: assume all rv's in the array are indpendent.
2. Centering: assume $\mathbb{E}X_{i,j} = 0$ for all i, j.
3. Variances converge: assume $\sum_{j=1}^{n} \mathbb{E}[X_{n,j}^2] \to \sigma^2 > 0$ as $n \to \infty$.
4. No single variance is too large: For all $\epsilon > 0$,

$$\lim_{n\to\infty} \sum_{j=1}^{n} \mathbb{E}[X_{n,j}^2 | |X_{n,j}| > \epsilon] = 0$$

Then $S_n \xrightarrow{D} N(0, \sigma^2)$ as $n \to \infty$.

Examples

1. Show that the theorem for triangular arrays is a generalization of the standard CLT.
2. Show that if we have independent rv's with uniformly bounded means and variances then a CLT holds.

Gaussian Probabilities

Let $Z \sim N(0, 1)$.
Then

1. $\Pr[-1 \leq Z \leq 1] \approx .68$ (within 1 standard deviation)
2. $\Pr[-2 \leq Z \leq 2] \approx .95$ (within 2 standard deviations)
3. $\Pr[-3 \leq Z \leq 3] \approx .99$ (within 3 standard deviations)

Other probabilities can be estimated using the fact that Z is symmetric, and by scaling and centering non-standard Normals.