March 27, 2020

Dr. SUSHIL KUMAR AZAD

In this lecture, we shall study one of the most remarkable results in probability theory, viz., the Central limit theorem (CLT).

The Central limit theorem provides a simple method for computing approximate probabilities for sums of independent random variables.

Pierre-Simon Laplace, a theoretical astrologer and mathematician, is best known to statisticians for the Central limit theorem which he arrived at non-rigorously in 1810.

This theorem also explains the remarkable fact that "the empirical frequencies of so many natural "populations" exhibit a normal curve".

SUSHIL KUMAR AZAD

We already know that a sequence of Binomial r.v.'s $\{X_n\}$ converges to a normal variate as $n \to \infty$. The CLT shows that this result is true for an arbitrary sequence of r.v.'s under certain conditions.

We shall first study the convergence of sequence of r.v.'s.

Definition 1. A sequence of r.v.'s $\{X_n\}$ defined over a sample space, $\mathcal{G}$, is said to converge to a r.v. X, also defined on $\mathcal{G}$, if $\{X_n(e)\}$ converges to $X(e) < \infty$ as $n \to \infty$ $\forall$ $e \in \mathcal{G}$. (In this case $\{X_n\}$ is said to converge to X everywhere).

Remark

There are a number of different ways i.e. different meanings attached to the statement

$$\lim_{n\to\infty} X_n = X.$$

Definition 2. (Convergence in Probability)

A sequence of r.v.'s $\{X_n\}$ is said to converge in probability (or stochastically) to a r.v. X, denoted as $X_n \xrightarrow{P} X$, if, for every $\varepsilon > 0$, as $n \to \infty$,

$$P[\,|X_n - X| \geq \varepsilon\,] \to 0$$

or (equivalently)

$$P[\,|X_n - X| < \varepsilon\,] \to 1.$$

The convergence in Probability means that for large n, the difference between X_n and X is likely to be small with large probability.

SUSHIL KUMAR AZAD

Note. The definition 2 does not imply that $|X_n(c) - X(c)|$ is small $\forall\, c \in \mathcal{C}$, however large n may be, only the probability for this is large.

Definition 3. (Convergence in Distribution)

Let $F_n(\cdot)$ be the cdf of a r.v. X_n and $F(\cdot)$ be that of the r.v. X. Then the sequence $\{X_n\}$ is said to converge to X in distribution (or in law) if $F_n(x) \to F(x)$ $\forall\, x \in A$, where A is the set of points of continuity of $F(x)$, and is denoted as

$$X_n \xrightarrow{L} X.$$

Theorem 1 (CLT): Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables, each with mean μ and variance, σ^2. Then the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal as $n \to \infty$.
In other words, for $-\infty < a < \infty$, we have

$$P\left\{ \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a \right\} \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\, dx,$$

as $n \to \infty$.

The Central limit theorem in the case of equal components is also stated as follows:

Theorem 2 (Lindeberg-Levy Central limit theorem)

Let $\{X_k\}$, $k = 1, 2, \cdots, n$ be a sequence of independent, and identically distributed r.v.'s with finite mean, μ and finite variance, σ^2. Let

$$Y_n = X_1 + X_2 + \cdots + X_n.$$

Then the sequence $\{Y_n\}$ converges in distribution to a r.v. Y which is normally distributed with mean, $n\mu$ and variance, $n\sigma^2$, i.e.

$$Y_n \xrightarrow{L} Y \text{ as } n \to \infty,$$

and we say that Y_n is Asymptotically Normally Distributed

SUNIL KUMAR AZAD

For the proof of CLT, we shall use the following result –

<u>Lemma 1.</u> Let $X_1, X_2, \dots$ be a sequence of r.v.'s having cdf's, $F_n(t) \left(= P[X_n \le t]\right)$ and MGF's, $M_n(t) \left(= E[e^{tX_n}]\right)$. Let X be a r.v. with cdf. $F(t)$ and MGF, $M(t)$.

If $M_n(t) \longrightarrow M(t)$ for all t,
then $F_n(t) \longrightarrow F(t)$ for all t at which $F(t)$ is continuous, i.e.

$$\boxed{M_n(t) \to M(t) \ \forall\, t \implies F_n(t) \xrightarrow{L} F(t)}$$

Note. We know that if $X \sim N(0,1)$, then

$$M(t) = e^{t^2/2}.$$

Thus, by lemma 1, if

$$M_n(t) \to e^{t^2/2} \text{ as } n \to \infty,$$

then, $F_n(t) \xrightarrow{L} \Phi(t)$ as $n \to \infty$,

where $\Phi(t)$, is the probability that the standard normal is less than t, i.e.,

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\, dx$$

We now present a heuristic proof (based on above lemma 1) of the central limit theorem.

Continued ...

Proof (CLT). Let us assume at first that
$\mu = 0$ and $\sigma^2 = 1$.

We shall prove the CLT under the assumption that the M.G.F. of the X_n, $M_n(t)$ exists and is finite.

Now, the M.G.F. of $\left(\frac{X_i}{\sqrt{n}}\right)$ is given by

$$E\left[\exp\left\{\frac{tX_i}{\sqrt{n}}\right\}\right]$$

$$= M_i\left(\frac{t}{\sqrt{n}}\right),$$

and thus the M.G.F. of $\left(\sum_{i=1}^{n} X_i / \sqrt{n}\right)$ is given by

$$E\left[\exp\left\{\frac{t(X_1+\cdots+X_n)}{\sqrt{n}}\right\}\right] = E\left[e^{tX_1/\sqrt{n}} \cdot e^{tX_2/\sqrt{n}} \cdots e^{tX_n/\sqrt{n}}\right]$$

$$= E\left(e^{tX_1/\sqrt{n}}\right) \times \cdots \times E\left(e^{tX_n/\sqrt{n}}\right)$$

$$= M_1\left(t/\sqrt{n}\right) \times \cdots \times M_n\left(t/\sqrt{n}\right)$$

$$= \left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n \quad \text{---(i)}$$

where $M\left(\frac{t}{\sqrt{n}}\right)$ is the common M.G.F. of $(X_i/\sqrt{n})$'s

Let $L(t) = \log(M(t))$.

We have $L(0) = 0$,

$$L'(0) = \frac{M'(0)}{M(0)} = \mu = 0,$$

$$L''(0) = \frac{M(0)\,M''(0) - [M'(0)]^2}{[M(0)]^2}$$

$$= E(X^2) = 1$$

Continued ..

SUSHIL KUMAR AZAD

Now, $\lim\limits_{n\to\infty} \dfrac{L(t/\sqrt{n})}{n^{-1}}$ $\left(\frac{0}{0}\text{ form}\right)$

$$= \lim_{n\to\infty} \left[\frac{-L'(t/\sqrt{n}) \cdot n^{-3/2}\, t}{-2\, n^{-2}}\right] \quad \text{(by L'Hôspital's rule)}$$

$$= \lim_{n\to\infty} \left[\frac{L'(t/\sqrt{n})\, t}{2\, n^{-1/2}}\right] \quad \left(\frac{0}{0}\text{ form}\right)$$

$$= \lim_{n\to\infty} \left[\frac{-L''(t/\sqrt{n}) \cdot n^{-3/2}\, t^2}{-2\, n^{-3/2}}\right] \quad \text{(again by L'Hôspital's rule)}$$

$$= \lim_{n\to\infty} \left[L''(t/\sqrt{n})\, \frac{t^2}{2}\right]$$

$$= \frac{t^2}{2}$$

$$\Rightarrow \lim_{n\to\infty} n\, L(t/\sqrt{n}) = t^2/2$$

$$\Rightarrow \lim_{n\to\infty} [M(t/\sqrt{n})]^n = e^{t^2/2}$$

that is, $[M(t/\sqrt{n})]^n \to e^{t^2/2}$ as $n\to\infty$

or M.G.F. of $\left(\sum_{i=1}^{n} x_i/\sqrt{n}\right) \to e^{t^2/2}$ as $n \to \infty$ from i)

$\therefore$ By lemma 1, $\left(\dfrac{\sum_{i=1}^{n} x_i}{\sqrt{n}}\right) \xrightarrow{L}$ a standard normal variate as $n\to\infty$

SUNIL KUMAR AZAD

Continued ...

Thus, the CLT is proved when $\mu = 0$ and $\sigma = 1$.

The general case, now, can be shown by considering the standardized r.v.'s

$$Z_i = \frac{X_i - \mu}{\sigma},$$

and applying the above proof, since $E(Z_i) = 0$, $var(Z_i) = 1$, and

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

$$= \frac{(X_1 - \mu) + (X_2 - \mu) + \cdots + (X_n - \mu)}{\sigma\sqrt{n}}$$

$$= \frac{Z_1 + Z_2 + \cdots + Z_n}{\sqrt{n}}$$

$$= \frac{\left(\sum_{i=1}^{n} Z_i\right)}{\sqrt{n}}. \quad //$$

SUSHIL KUMAR AZAD

<u>Remark</u> (1) The first version of the CLT was proved by DeMoivre around 1733 for the Bernoulli r.v.'s with $p = \frac{1}{2}$.

(2) Laplace also discovered the more general form (for arbitrary p) as given in Theorem 1, but his proof was not completely rigorous.

(3) A truly rigorous proof of CLT was first presented by Russian mathematician <u>Liapounoff</u> in the period 1901–1902. //