Central Limit Theorem using Characteristic functions

RongXi Guo

MAT 477

January 20, 2014

Introduction–study a random variable

Let Ω with measure m , $m(\Omega) = 1$ and $\mathfrak{F}(\Omega)$ measurable functions.

To random variables $X \in \mathfrak{F}(\Omega)$ we associate distribution functions

$F(x) := P(X < x) := m(\xi \in \Omega : X(\xi) < x)$ and $f(x) = F'(x)$ is the

probability density function (shortly pdf). We assume exist finite:

Expected value of X (mean value) $\mu = E(X) = \int_R x dF(x) = \int_\Omega X(\xi) dm$

Variance and the standard deviation $\sigma^2 = V(X) = \int_R (x - \mu)^2 dF(x)$.

Convention: $P(...) := m(...)$ and for $\{A_j \subset R\}_{1 \le j \le n}$ and

$\{X_j \in \mathfrak{F}(\Omega)\}_j$ set $X_1 \in A_1, \cdots, X_n \in A_n := \{\xi \in \Omega : X_j(\xi) \in A_j, \forall j\}$

For our X_j distribution function, expected value and variance

are the same and $\{X_j\}_j$ are independent, identically distributed

(shortly iid), i.e. $P(X_1 \in A_1, \cdots, X_n \in A_n) = \Pi_{j=1}^{n} P(X_j \in A_j)$.

For iid $\{X_j\}_j$ set $S_n := \frac{\sum_1^n X_i}{n}$ and also

$$E(S_n) = E\left(\frac{\sum_1^n X_i}{n}\right) = \frac{\sum_1^n E(X_i)}{n} = \frac{n \cdot \mu}{n} = \mu$$

$$V(S_n) = V\left(\frac{\sum_1^n X_i}{n}\right) = \frac{V\left(\sum_1^n X_i\right)}{n^2} = \frac{n \cdot V(X)}{n^2} = \frac{V(X)}{n} = \frac{\sigma^2}{n} .$$

Law of Large Numb: $\lim_{n\to\infty} P(|S_n - \mu| > \epsilon) = 0, \forall\epsilon > 0$

Proof: $\frac{\sigma^2}{n} = V(S_n) = \int_R (x-\mu)^2\, dF(x) \geq \int_{|x-\mu|>\epsilon} (x-\mu)^2\, dF(x)$

$\geq \epsilon^2 \cdot P(|S_n - \mu| > \epsilon) \Rightarrow P(|S_n - \mu| > \epsilon) \leq \delta$ *for* $n > \frac{\sigma^2}{\epsilon^2 \cdot \delta}$.

Def: $X_n \xrightarrow{d} X$, i.e. converge in the sense of distributions means

$\forall$ bounded and continuous function f : $\int_R f dF_n \to \int_R f dF$.

Central Limit Theorem (shortly CLT): $\frac{(S_n-\mu)\sqrt{n}}{\sigma} \xrightarrow{d} N(0,1)$, where

$S_n = \frac{\sum_1^n X_i}{n}$ and $N(0,1)$ is the rv with pdf $\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$ of Gauss distribution

Next, note that $N(0,1)$ **has expected value** $\int_R x dF(x) = 0$ **and**

variance $\int_R x^2 dF(x) = 1$. Also, $\{X_j\}_j$ being iid's of course (page 3)

$E(S_n) = \mu$, $V(S_n) = \frac{\sigma^2}{n}$. To prove the theorem we'll use the

characteristic functions $\varphi(t) = E(e^{itX}) = \int_R e^{itx} dF(x)$, shortly cfs

Note: rvs always admit cfs; $\varphi'(0) = i\mu$ and $\mu = 0 \Rightarrow \varphi''(0) = -\sigma^2$.

Also $\varphi(0) = E(1) = 1$, $|\varphi(t)| = \left|\int_R e^{itx} dF(x)\right| \leq \int_R \left|e^{itx}\right| dF(x) = 1$.

Fact1: $F(b) - F(a) = \frac{1}{2\pi} \lim_{x \to \infty} \int_{-x}^{x} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)\, dt$.

Easy if $\exists\ F'(x)$: $f(x) = \frac{1}{2\pi} \int_R e^{-itx} \varphi(t)\, dt$.

Properties of characteristic function

$$\varphi_{X_1+X_2}(t) = E(e^{it(X_1+X_2)}) = E(e^{itX_1}) \cdot E(e^{itX_2}) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t)$$

$$\varphi_{aX+b}(t) = E(e^{it(aX+b)}) = e^{itb} \cdot E\left(e^{i(at)X}\right) = e^{itb} \cdot \varphi_X(at)$$

and also the uniform continuity of cf with $\mu = 0$ and $\sigma = 1$:

$$|\varphi(t+h) - \varphi(t)| = \left|E(e^{i(t+h)X} - e^{itX})\right| \leq E(\left|e^{ihX} - 1\right|) \to 0$$

Characteristic function for Gauss distribution is $e^{-\frac{t^2}{2}}$, page 14.

Convergence of F implies convergence of cfs

Proposition : $X_n \xrightarrow{d} X \Leftrightarrow \varphi_n(x) \to \varphi(x)\ \forall x \in R$.

Proof. "$\Rightarrow$": e^{itx} is bounded and continuous and $X_n \xrightarrow{d} X$

imply $\int_R e^{itx} dF_n \to \int_R e^{itx} dF$. To show "$\Leftarrow$" (see page 12)

we need to prove first a so called 'tightness' of our rvs.

Tightness of a family of Random Variables.

Def: a family of rvs X_n is tight when

$\forall \epsilon > 0\ \exists\ M$ such that $P(|X_n| > M) < \epsilon$ for all n .

Claim: convergence of cfs implies tightness of rvs .

Proof of 1st step : we show that for any distribution $X := X_n$,

$\forall \epsilon > 0 \ \exists \ M \ , \ P\left(|X| > M\right) < \frac{\epsilon}{2}$.

Indeed, every cf has a value of 1 at 0 (page 5) and is continuous

$\Rightarrow \forall \ \epsilon > 0 \ \exists \ \delta > 0$ such that $\forall \ |t| < \delta \ , \ |1 - \varphi\left(t\right)| < \frac{\epsilon}{4}$

$\Rightarrow \int_{-\delta}^{\delta} |1 - \varphi\left(t\right)| \, dt < 2\delta \cdot \frac{\epsilon}{4} = \frac{\epsilon \cdot \delta}{2}$

$\Rightarrow \delta^{-1} \int_{-\delta}^{\delta} |1 - \varphi\left(t\right)| \, dt < \frac{\epsilon}{2}$. On the other hand, for some large M

$\delta^{-1} \int_{-\delta}^{\delta} |1 - \varphi\left(t\right)| \, dt$ is an upper bound on $P(|X| \geq M)$:

$\delta^{-1}\int_{-\delta}^{\delta}(1-\varphi(t))\,dt = \delta^{-1}\int_{-\delta}^{\delta}\left(1-E\left(e^{itX}\right)\right)dt$

$= \delta^{-1}\left(2\delta - \int_R \left(\frac{\sin(\delta x)}{x} - \frac{\sin(-\delta x)}{x}\right) dF(x)\right)$

$= 2\left(1-\int_R \frac{\sin(\delta x)}{\delta x} dF(x)\right)$. Replacing 1 by $\int_R 1 dF(x)$

we have $2\left(1-\int_R \frac{\sin(\delta x)}{\delta x} dF(x)\right) = 2\int_R \left(1-\frac{\sin(\delta x)}{\delta x}\right) dF(x)$.

$2\int_R \left(1-\frac{\sin(\delta x)}{\delta x}\right) dF(x) \geq 2\int_{|x|\geq\frac{2}{\delta}} \left(1-\frac{\sin(\delta x)}{\delta x}\right) dF(x) \ \geq$

$\int_{|x|\geq\frac{2}{\delta}} 1 dF(x) = P\left(|X|\geq\frac{2}{\delta}\right) \ \Rightarrow\ \delta^{-1}\int_{-\delta}^{\delta}|1-\varphi(t)|\,dt \geq P\left(|X|\geq\frac{2}{\delta}\right)$.

Together with above $\frac{\epsilon}{2} > \delta^{-1}\int_{-\delta}^{\delta}|1-\varphi(t)|\,dt \ \geq P\left(|X|\geq\frac{2}{\delta}\right)$. $\square$

Step 2 : convergence of cfs implies tightness in its rvs

$\varphi_n(x) \to \varphi(x)$ means $\forall\ \epsilon > 0\ ,\ x \in R\ \exists$ natural number N s. th.

$\forall\ n > N$ holds $|\varphi_n(x) - \varphi(x)| < \frac{\epsilon}{4} \quad \Rightarrow\ \forall\ \epsilon\ ,\ \delta\ \ \exists\ N$ such that

$\forall\ n \geq N$ we have $\delta^{-1} \int_{-\delta}^{\delta} |\varphi_n(t) - \varphi(t)|\, dt < \frac{\epsilon}{2}$ (**fact from analysis**).

Also, (page 9) we may choose δ to satisfy $\delta^{-1} \int_{-\delta}^{\delta} |1 - \varphi(t)|\, dt < \frac{\epsilon}{2}$

$\forall\ n \geq N$ we have $P\left(|X_n| \geq \frac{2}{\delta}\right) \leq \delta^{-1} \int_{-\delta}^{\delta} |1 - \varphi_n(t)|\, dt$

$$\leq \delta^{-1}\left(\int_{-\delta}^{\delta}|1-\varphi(t)|\,dt+\int_{-\delta}^{\delta}|\varphi_n(t)-\varphi(t)|\,dt\right)<\epsilon\ .$$

Also, for n smaller than $N \ \exists\ \delta_n$ such that

$$P\left(|X_n| \geq \frac{2}{\delta_n}\right) \leq \delta_n^{-1}\int_{-\delta_n}^{\delta_n}|1-\varphi_n(t)|\,dt<\epsilon$$

choose $\delta_{min} := \min\{\delta_1, \delta_2, \cdots, \delta_n, \delta\}$.

we have then that $P\left(|X_n| \geq \frac{2}{\delta_{min}}\right) < \epsilon$ for any n

$\Rightarrow$ rvs with convergent cfs are tight, the claim is proved. $\square$

Proof of $\varphi_n(x) \to \varphi(x)\ \forall\, x$ implies $X_n \xrightarrow{d} X$ using

Fact 2. Tightness of rvs implies compactness in the sense of convergence of distributions ("**Prokhorov's Theorem**").

Proof of "$\Leftarrow$" from page 7 : Pick any convergent, say to F_1 , subsequence $\{F_{1n}\}_n$ of distributions. Say $\{\varphi_{1n}\}_n$ are their cfs.

$\varphi_n(x) \to \varphi(x)\ \forall\, x$ implies convergence of all $\{\varphi_{1n}\}_n$ to the same φ and proved "$\Rightarrow$" on page 7 implies that φ is the cf for any F_1

$\Rightarrow$ exists unique $F_1 =: F$ and, using **Fact 2.**, $\Rightarrow$ $X_n \xrightarrow{d} X$, i.e.

$X_n \xrightarrow{d} X \Leftrightarrow \varphi_n(x) \to \varphi(x)\ \forall\, x$ is proved. □

Conclusion of the proof of Central Limit Theorem

For a series of iid X_i, let $Y_n = \frac{\sum_1^n X_i - n\mu}{\sigma\sqrt{n}}$

$\varphi_{Y_n}(t) = \varphi_{\frac{\sum_1^n X_i - n\mu}{\sigma\sqrt{n}}}(t) = \varphi_{\sum_1^n X_i - n\mu}(\frac{t}{\sigma\sqrt{n}}) =: \varphi^n(\frac{t}{\sigma\sqrt{n}})$. Let $s := \frac{t}{\sigma\sqrt{n}}$,

as $n \to \infty$ $s \to 0$. Recall: $\varphi'(0) = i\mu = 0$, $\varphi''(0) = -\sigma^2$, see page 5 .

From Taylor expansion: $\varphi(0) + s \cdot \varphi'(0) + \frac{s^2}{2} \cdot \varphi''(0) - \varphi(s) = o(s^2) \Rightarrow$

$\lim_{n\to\infty}\{\varphi_{Y_n}(t) = \left(1 - \frac{\sigma^2 + o(1)}{2} \cdot (\frac{t}{\sigma\sqrt{n}})^2\right)^n\} = \lim_{n\to\infty}\left(1 - \frac{\sigma^2}{2}(\frac{t}{\sigma\sqrt{n}})^2\right)^n$

$= e^{-\frac{t^2}{2}} \Rightarrow$ the limit of the cfs is the cf of a Gauss distribution

$\Rightarrow \quad \frac{(S_n - \mu)\sqrt{n}}{\sigma} = Y_n \xrightarrow{d} N(0,1)$, as required. $\square$

Appendix. cf of Normal distribution, calculation:

$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.

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

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2 + itx - \frac{1}{2}(it)^2 + \frac{1}{2}(it)^2} dx$

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

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

$y = x - it \Rightarrow \frac{dy}{dx} = 1 \Rightarrow$

$\varphi(t) = e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy = e^{-\frac{t^2}{2}}$.

Abbreviations

rv : random variable

rvs : random variables

pdf : probability density function

iid : independent, identical distributed rvs

cf : characteristic function

cfs : characteristic functions