ERASMUS, BS course
Winter 2020/2021
FPM&H, December 10, part II

Subject. Central Limit Theorem and its application.

Introduction.

From the 1st part we know that if we have the sequence $(X_n)_{n\geq 1}$, where $n \geq 1$ $d(X_n) = d(X)$ for some X with second moment and (X_n) are i.i.d., then

$$\frac{1}{n}(X_1 + X_2 + \ldots + X_n) \xrightarrow{a.s.} EX = m.$$

What can happen if we consider the stronger type of convergent, namely with respect of distribution?

The Central Limit Theorem (CLT) gives the answer to that question.

Assumptions

Let X and (X_n) as above and in additional, $\sigma_X^2 = var(X) > 0$ (so $X \neq const$).

Let us take:

1) $Y_n \stackrel{df}{=} \frac{1}{n}(X_1 + X_2 + \dots + X_n)$ (the average)

Then

$$EY_n = \frac{1}{n} \cdot n\, EX = m,$$

$$var(Y_n) = \frac{1}{n^2} n\, var(X) = \frac{\sigma^2}{n} \quad (\sigma^2 = \sigma_X^2).$$

2) After standarization procedure of Y_n we get for $n \geq 2$ and by the above

$$Z_n \stackrel{df}{=} \frac{Y_n - EY_n}{\sqrt{var(Y_n)}} = \frac{Y_n - m}{\frac{\sigma}{\sqrt{n}}}$$

then C.L.T. says that:

$$Z_n \xrightarrow{d} \mathcal{N}, \text{ where } \mathcal{N} \in N(0,1)$$

So we have

$\forall t \in \mathbb{R}$

$$F_{Z_n}(t) = P(\{\omega \in \Omega : Z_n(\omega) < t\}) \longrightarrow$$

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

REMARKS.

1). For sufficient large $n \geq 2$ we have an approximation rule

$$F_{Z_n} \approx \Phi$$

2) We will check what this means for X:

We know from the above that

$$F_{Z_n}(t) = P\left(\{\omega\in\Omega : \frac{Y_n(\omega)-m}{\frac{\sigma}{\sqrt{n}}} < t\}\right) =$$

$\downarrow$

$$\Phi(t) = P\left(\{\omega\in\Omega : Y_n(\omega) < m + t\frac{\sigma}{\sqrt{n}}\}\right) =$$

$$= F_{Y_n}\left(m + \frac{t\sigma}{\sqrt{n}}\right), \text{ for } t\in\mathbb{R}.$$

So, by C.L.T. we can write

$$F_{Y_n}\left(m + \frac{t\sigma}{\sqrt{n}}\right) \approx \Phi(t) \quad (*)$$

Let $u = m + \frac{t\sigma}{\sqrt{n}} \Rightarrow t = \frac{u-m}{\frac{\sigma}{\sqrt{n}}}$, and $(*)$ means that

$$F_{Y_n}(u) = \Phi\left(\frac{u-m}{\frac{\sigma}{\sqrt{n}}}\right).$$

But

$$\Phi\left(\frac{u-m}{\frac{\sigma}{\sqrt{n}}}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{u-m}{\frac{\sigma}{\sqrt{n}}}} e^{-t^2/2}\,dt.$$

Let us apply the substitution rule in this integral:

$$t = \frac{u-m}{\frac{\sigma}{\sqrt{n}}} \qquad dt = \frac{\sqrt{n}}{\sigma}\, du \quad , \text{so}$$

$$\Phi\left(\frac{u-m}{\frac{\sigma}{\sqrt{n}}}\right) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{n}}{\sigma} \int_{-\infty}^{t} e^{-\left(\frac{u-m}{\frac{\sigma}{\sqrt{n}}}\right)^2 / 2} du$$

$$= \underbrace{\frac{1}{\sqrt{2\pi \frac{\sigma^2}{n}}} \int_{-\infty}^{t} e^{-\frac{(u-m)^2}{2\frac{\sigma^2}{n}}} du}_{\text{the cummulative dist. function corresponding to } N(m, \frac{\sigma^2}{n})}.$$

So, finally

$$F_{Y_n} \approx N\left(m, \frac{\sigma^2}{n}\right)$$

Let us consider the following situation:

for $n \geq 2$ $\quad a < b$, $\quad B_n \in B(n,p)$

$p \in (0,1)$

$$A_{a,b,n} = \{\omega \in \Omega : a \leq B_n(\omega) < b\}$$

and let us <u>try to calculate $P(A_{a,b,n})$</u>.

<u>We will get ~~[illegible]~~</u>: successively after averaging

$$P(A_{a,b,n}) = P\left(\left\{\omega \in \Omega : \frac{a}{n} \leq \frac{B_n(\omega)}{n} < \frac{b}{n}\right\}\right)$$

But $E\left(\frac{B_n}{n}\right) = p$, $Var\left(\frac{B_n}{n}\right) = \frac{pq}{n}$,

so after <u>standardization procedure we get</u>

$$= P\left(\left\{\omega \in \Omega : \frac{\frac{a}{n} - p}{\sqrt{\frac{pq}{n}}} \leq \frac{\frac{B_n(\omega)}{n} - p}{\sqrt{\frac{pq}{n}}} < \frac{\frac{b}{n} - p}{\sqrt{\frac{pq}{n}}}\right\}\right)$$

Finally, by C.L.T. we can write

$$P(A_{a,b,n}) \approx \Phi\left(\frac{\frac{b}{n}-p}{\sqrt{\frac{pq}{n}}}\right) - \Phi\left(\frac{\frac{a}{n}-p}{\sqrt{\frac{pq}{n}}}\right).$$

Remark.

The last formula is a first version of C.L.T - it is called de Moivre-Laplace Theorem (from XVIII c.)

Example.

Let $n = 100$, $a = 30$, $b = 70$, $p = 0,25$

Then

$$P(A_{a,b,n}) \cong \Phi(10,3923) - \Phi(1,1547)$$

$$\cong 1 - \Phi(1,1547) =$$

$$= 1 - 0,8943 = \underline{\underline{0,1057}}.$$

Example:

We will estimate

$$(*)\quad P(\{\omega\in\Omega : B_n(\omega)=k\}) \text{ for } n=100,\ k=30,\ p=0{,}25.$$

We note, that

$$(*) = P(\{\omega\in\Omega : 29{,}5 \leq B_n(\omega) < 30{,}5\})$$

$$\approx \Phi(1{,}2701) - \Phi(1{,}03923)$$

$$= \underline{\underline{0{,}0495}}.$$

On the other hand we have:

$$P(\{\omega\in\Omega : B_{100}(\omega)=30\}) = \binom{100}{30}(0{,}25)^{30}(0{,}75)^{70}$$

But:

$$\binom{100}{30} = 2{,}9372 \cdot 10^{25}$$

$$(0{,}25)^{30} = 8{,}6736 \cdot 10^{-19}$$

$$(0{,}75)^{70} = 1{,}7959 \cdot 10^{-9}$$

and finally

$$P(\{w \in \Omega : B_{100}(w) = 30\}) = \underline{0{,}045}.$$

If we compare this results we get

$$|0{,}0495 - 0{,}045| = \underline{\underline{0{,}004}}.$$

<u>SO IT WORKS !!!</u>