Clase 20, leer Capítulo 5. Limit Theorems, p. 263 a 284 del libro Introduction to Probability, 2nd ed. de Bertsekas, D.P. y Tsitsiklis, J.N.

Teorema Central de Límite

Sea $X_1, X_2, \ldots, X_n$ una m.a. de alguna distribución $f_X(x)$, en donde $E(X) = \mu$ y $Var(X) = \sigma^2 < \infty$

Si definimos $$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

$$\Rightarrow \lim_{n \to \infty} F_{Z_n}(z) = F_Z(z) \quad , \quad \forall z \in \mathbb{R}$$

en donde $Z \sim N(0,1)$

Ejemplo.

Invitamos a 64 personas a una fiesta. Hacemos sandwiches para los invitados. Creemos que

1: Cada invitado espera 0, 1 ó 2 sandwiches con probabilidades $1/4$, $1/2$, $1/4$.

2: El número de sandwiches que cada invitado come es independiente al de los demás invitados

¿Cuantos sandwiches necesitamos preparar para estar 95% seguros de que todos queden satisfechos?

Primero definimos X_i, el número de sandwiches que el invitado i-esimo come

$$\Rightarrow \quad \underset{i=1,\dots,60}{P(X_i = x)} = P(X = x) = \begin{cases} 1/4, & x = 0 \\ 1/2, & x = 1 \\ 1/4, & x = 2 \end{cases}$$

$$\mathbb{E}(X) = 1/4(0) + 1/2(1) + 1/4(2) = 1 = \mu$$

$$\mathbb{E}(X^2) = 1/4(0^2) + 1/2(1^2) + 1/4(4) = 3/2$$

$$Var(X) = 3/2 - 2/2 = 1/2 = \sigma^2$$

así el numero de sandwiches que comen los invitados en la fiesta está dado por

$$Y = \sum_{i=1}^{n} X_i \sim ¿?$$

necesitamos determinar $\underline{\underline{y}}$ tal que

$$P(Y \leq y) \geq 0.95$$

$$\Rightarrow \quad P(\overline{X}_n \leq y/n) = P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq \frac{y/n - \mu}{\sigma/\sqrt{n}}\right)$$

TCL

$$\approx P\left(Z \leq 8\sqrt{2}\left(\frac{y}{64} - 1\right)\right) \geq 0.95$$

$$= P\left(Z \leq \frac{\sqrt{2}}{8}(y-64)\right) \geq 0.95$$

$$= F_Z\left(\frac{\sqrt{2}}{8}(y-64)\right) \geq 0.95$$

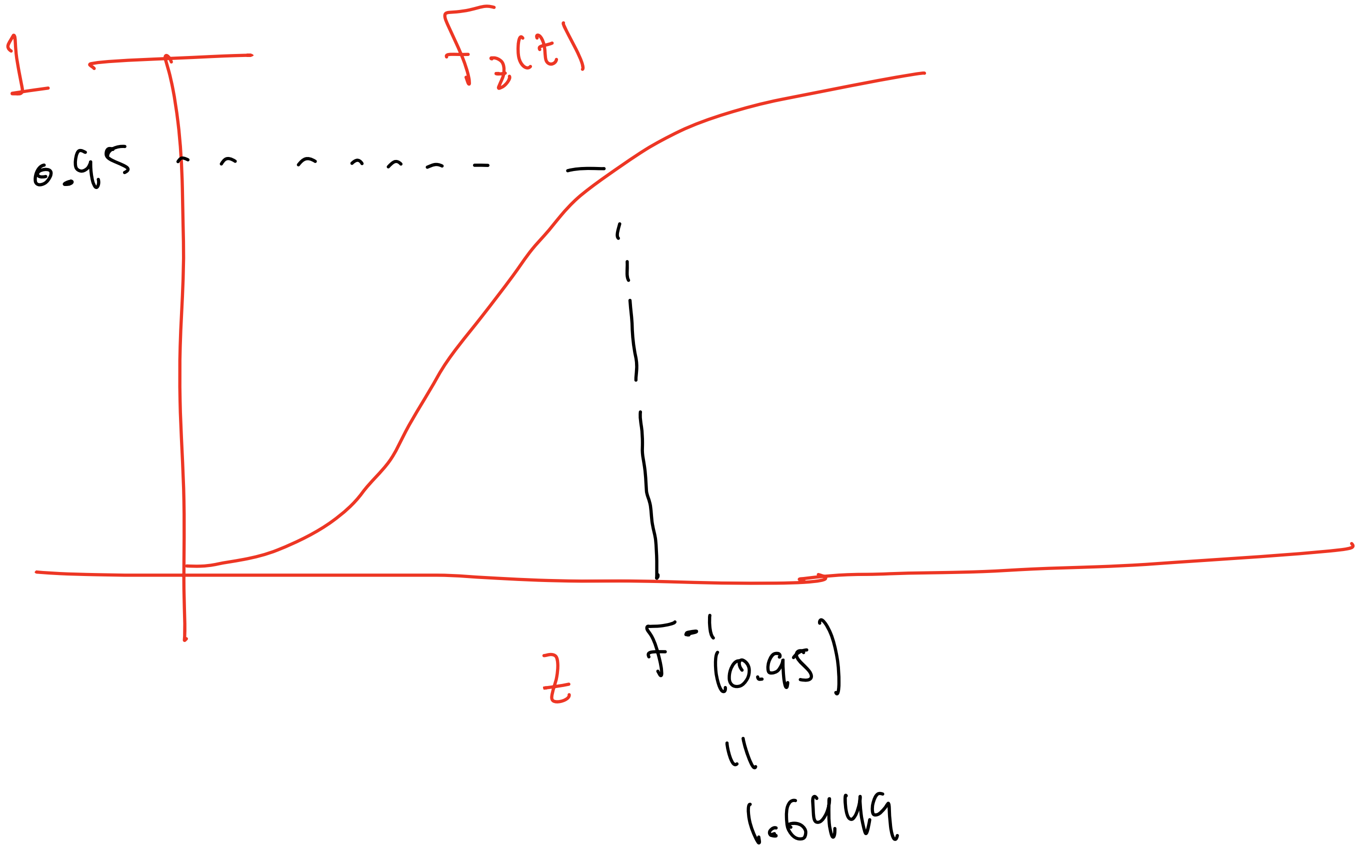
$$\Rightarrow \quad \frac{\sqrt{2}}{8}(y-64) \geq F_Z^{-1}(0.95)$$

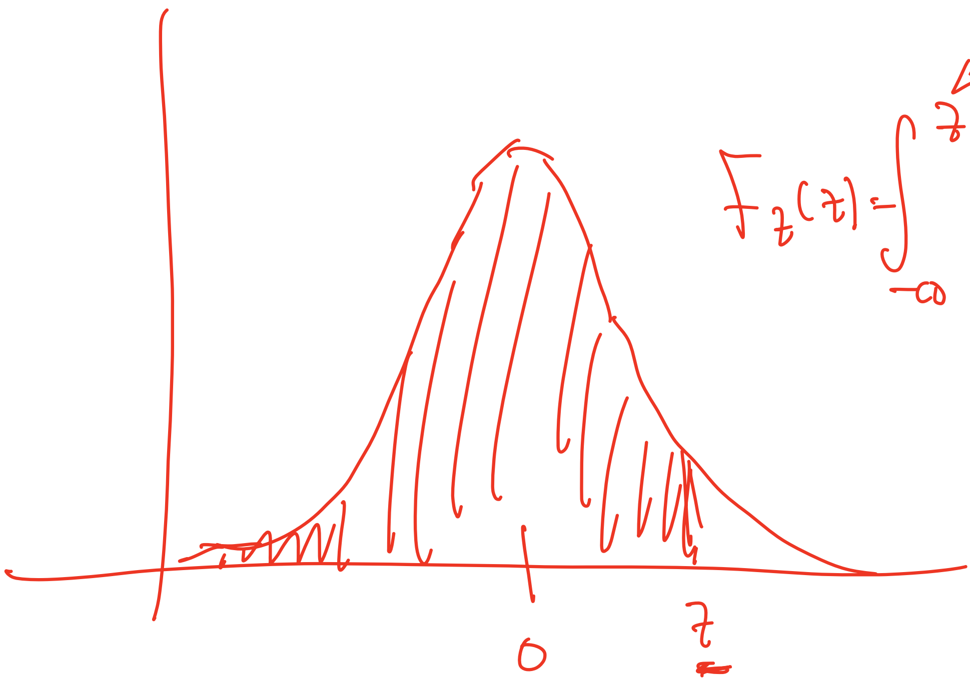
$$F_Z^{-1}(0.95) = \text{qnorm}(0.95, 0, 1)$$

$$\approx 1.645$$

$$y \geq \frac{8}{\sqrt{2}} F_Z^{-1}(0.95) + 64 = 73.3$$

$$\boxed{y = 74}$$





$$F_Z(z) = \int_{-\infty}^{z} f_Z(y)\,dy \geq 0.95$$

Ejemplo.

Se quiere determinar el tamaño necesario n que tiene que tener una m.a. $X_1, \dots, X_n$ con $X_i \sim \text{Bernoulli}(p)$ para asegurar que

$$P(|\overline{X}_n - p| \leq \varepsilon) \geq 0.95$$

en donde p es la proporción de personas que votarían por cierto candidato a un puesto de elección popular.

¿Qué tenemos?

$$\mathbb{E}(X) = p$$
$$Var(X) = p(1-p)$$

$$P(|\overline{X}_n - p| \leq \varepsilon) \geq 0.95 \Leftrightarrow 1 - P(|\overline{X}_n - p| \geq \varepsilon) \geq 0.95$$

$$P(|\overline{X}_n - p| \geq \varepsilon) \leq 0.05$$

Por Chebyshev tenemos la desigualdad

$$P(|\overline{X}_n - p| \geq \varepsilon) \leq \frac{Var(\overline{X}_n)}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \approx 0.05$$

$$\Rightarrow n \geq \frac{1}{4(0.05)\,\varepsilon^2}$$

¿Cómo sería usa el TCL?

$$P(|\overline{X}_n - p| \leq \varepsilon) \geq 0.95$$

$$\Rightarrow P(-\varepsilon \leq \overline{X}_n - p \leq \varepsilon) \geq 0.95$$

$\Leftrightarrow$

$$P\left(-\frac{\sqrt{n}\,\varepsilon}{\sigma} \leq Z \leq \frac{\sqrt{n}\,\varepsilon}{\sigma}\right) \geq 0.95$$

$\Leftrightarrow$ TCL

$$F_Z\left(\frac{\sqrt{n}\,\varepsilon}{\sigma}\right) - F_Z\left(-\frac{\sqrt{n}\,\varepsilon}{\sigma}\right) \geq 0.95$$

$$F_Z\left(\frac{\sqrt{n}\,\varepsilon}{\sigma}\right) - \left(1 - F_Z\left(\frac{\sqrt{n}\,\varepsilon}{\sigma}\right)\right) \geq 0.95$$

$$2F_Z\left(\frac{\sqrt{n}\,\varepsilon}{\sigma}\right) - 1 \geq 0.95$$

$$F_Z\left(\frac{\sqrt{n}\,\varepsilon}{\sigma}\right) \geq 0.05/2$$

$$\Leftrightarrow \quad n \geq \frac{\sigma^2}{\varepsilon^2} F^{-1}(0.05/2)^2$$

$$n \geq \frac{1}{4\varepsilon^2} F^{-1}(0.05/2)^2$$

Chebyshev $\quad n \geq \dfrac{1}{4\varepsilon^2\, 0.05}$

TCL $\quad n \geq \dfrac{1}{4\varepsilon^2} F^{-1}(0.05/2)^2$

	ε		
	0.1	0.05	0.01
Chebyshev	500	2000	50,000
TCL	97	385	9604