Central Limit Theorem (CLT)

The statement of CLT

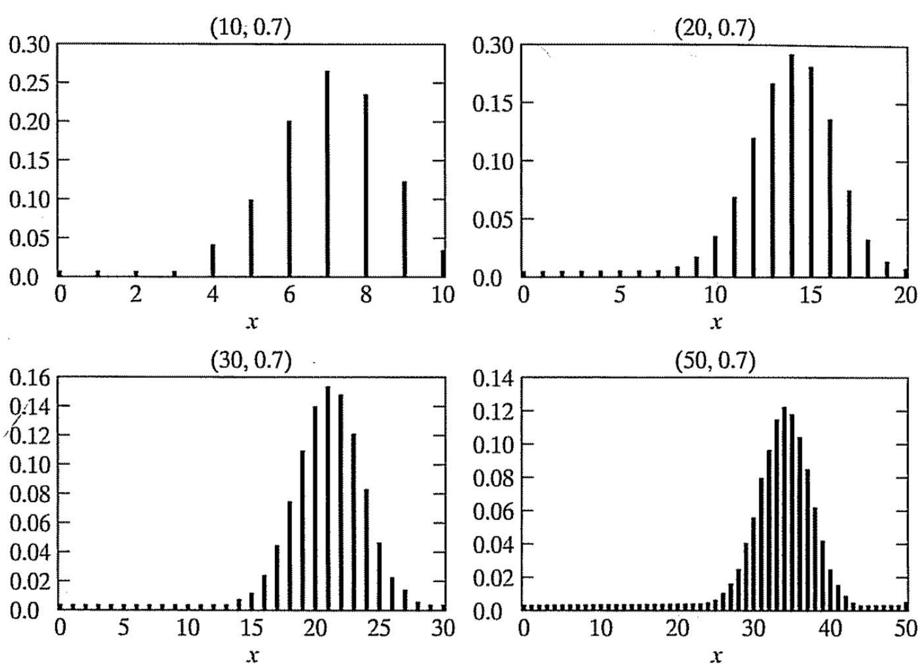
> If $X_1, \cdots, X_n$ are *i.i.d.* (independent and identically distributed) random variables (of any distribution) with mean μ and variance σ^2, and let $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$, then
>
> $$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$
>
> has a probability distribution that converges to $N(0,1)$ as $n \to \infty$.

- We say: Z has an asymptotic standard normal distribution. (asymptotic: $n \to \infty$)
- The statement can also be

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

  has a distribution that converges to $N(0,1)$ as $n \to \infty$.
- We can also say: $(X_1 + \cdots + X_n) \sim N(n\mu, n\sigma^2)$ as $n \to \infty$. Namely, it is the "sum" that is important (not "average").

Example: Bin(n,p), $n = 10 \to 20 \to 30 \to 50$, $p = 0.7$ fixed



- $\text{Bin}(n, p)$ can be seen as the distribution of the sum of n $\text{Ber}(p)$ random variables.
- The plots show that, as $n \uparrow$ $(10 \to 20 \to 30 \to 50)$, the distribution is more and more like a normal distribution.
- We see: $\text{Bin}(n, p) \to \text{Normal}(np, np(1-p))$ as $n \to \infty$ (with p fixed).

Example: "Triangle" distribution

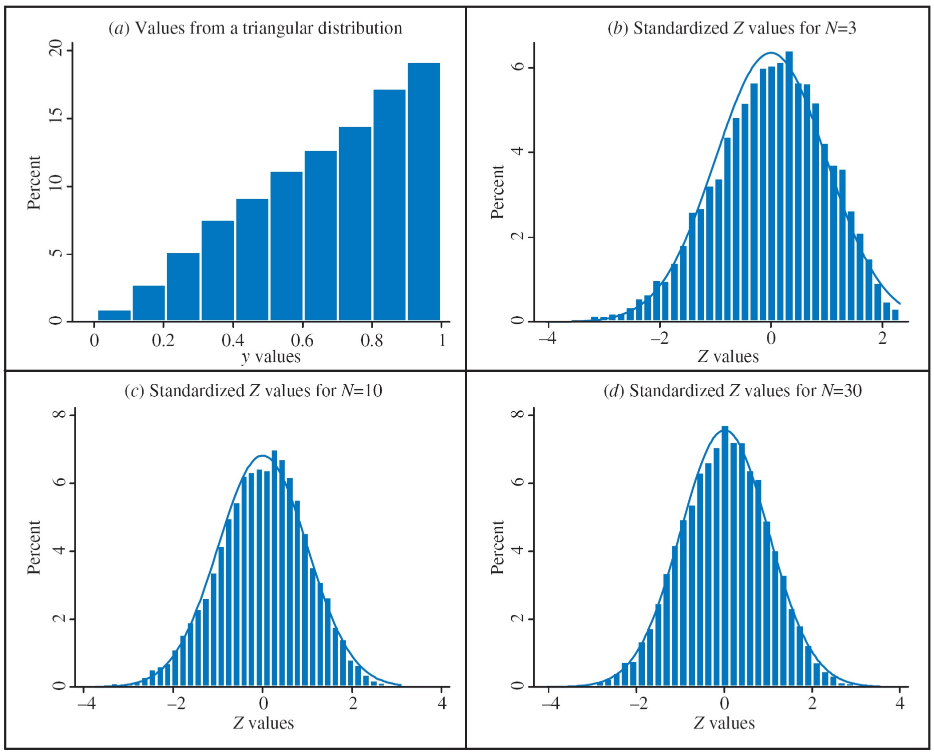


FIGURE **C.3** Central limit theorem.

- Let X_i follow a "triangle" distribution. Its mean is $\mu = 2/3$ and its variance is $\sigma^2 = 1/18$.
- Denote the sample mean as $\bar{X} = \sum_{i=1}^{n} X_i/n$. Its standardized version,

$$Z = \frac{\bar{X} - 2/3}{\sqrt{\frac{1/18}{n}}},$$

has a asymptotic standard normal distribution $(n = 3 \to 10 \to 30)$.