The Central Limit Theorem

- **Setup:** Let $X_1, X_2, X_3 \dots$ be **independent, identically distributed (iid)** random variables, i.e., r.v.'s that (i) are mutually independent, and (ii) have the same distribution with finite expectation and variance given by

$$\mu = E(X_i), \quad \sigma^2 = \mathrm{Var}(X_i),$$

  Let

$$S_n = \sum_{i=1}^{n} X_i.$$

- **Central Limit Theorem (CLT):**
  - **Normal approximation form:** For large n,

$$\boxed{S_n \text{ is approximately normal } N(n\mu, n\sigma^2)}$$

    or, equivalently,

$$\boxed{\frac{S_n - n\mu}{\sigma\sqrt{n}} \text{ is approximately standard normal } N(0,1)}$$

  - **Limit form:** For all z,

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right) = \Phi(z).$$

  - **Special case of normal r.v.'s:** In the special case when the X_i's are already normal with mean μ and σ^2, we know that the **exact** distribution of S_n is $N(n\mu, n\sigma^2)$, by the formula for the distribution of a sum of independent normals.

    Thus, the CLT can be interpreted as saying that, when n is large, then *a sum of n independent r.v.'s with an arbitrary common distribution behaves as if the individual variables were normal with the same mean and variance.*
  - **Normal approximation to the binomial distribution:** Another type of normal approximation came up in Chapter 5: A binomial distribution with parameters n and p is approximately normal $N(np, np(1-p))$ provided n is large and p is not too close to 0 or to 1. This result can be obtained as a special case of the CLT, by representing the number of successes in n S/F trials as a sum $S_n = \sum_{i=1}^{n} X_i$, where X_i is the indicator random variable of the event "success at trial i", and has mean $\mu = E(X_i) = p$ and variance $\sigma^2 = \mathrm{Var}(X_i) = p(1-p)$.

- **Weak Law of Large Numbers:** For any $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \le \epsilon\right) = 1.$$

- **Probabilistic inequalities:**
  - **Markov inequality:** Let X be a *nonnegative* (i.e., $X \ge 0$) r.v. with mean μ. Then, for any $a > 0$,

$$P(X \ge a) \le \frac{\mu}{a}.$$

  - **Chebychev inequality:** Let X be a r.v. with mean μ and variance σ^2. Then, for any $k > 0$,

$$P(|X - \mu| \ge k) \le \frac{\sigma^2}{k^2}.$$