Breaking down the theorem

Central Limit Theorem

Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables, with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma^2}$. Let $\boldsymbol{Y_n} = \frac{\boldsymbol{X_1+X_2+\cdots+X_n-n\mu}}{\boldsymbol{\sigma\sqrt{n}}}$

As $\boldsymbol{n} \to \infty$, the CDF of $\boldsymbol{Y_n}$ converges to the CDF of $\mathcal{N}(\mathbf{0}, \mathbf{1})$

A problem

What's the probability that $\mathrm{X} = 950$? (exactly)

True value, we can get with binomial:

$\binom{1000}{950} \cdot (.95)^{950} \cdot (.05)^{50} \approx .05779$

What does the CLT say?

Approximating a continuous distribution

You buy lightbulbs that burn out according to an exponential distribution with parameter of $\lambda = 1.8$ lightbulbs per year.

You buy a 10 pack of (independent) light bulbs. What is the probability that your 10-pack lasts at least 5 years?

Let X_i be the time it takes for lightbulb i to burn out.

Let X be the total time. Estimate $\mathbb{P}(X \geq 5)$.

Confidence Intervals

A "confidence interval" tells you the probability (how confident you should be) that your random variable fell in a certain range (interval)

Usually "close to its expected value"

$$\mathbb{P}(|X - \mu| > \varepsilon) \leq \delta$$

If your RV has expectation equal to the value you're searching for (like our polling example) you get a probability of being "close enough" to the target value.