

Probability Theory

Textbook: *Introduction to Probability* by Blitzstein and Hwang

Previous Lecture

- ◆ Cauchy-Schwarz Inequality, $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$.
- ◆ Jensen's Inequality: convex $E(g(X)) \geq g(E(X))$; concave $E(g(X)) \leq g(E(X))$.
- ◆ Bounds on tail probabilities.
 - ◆ Markov: $P(|X| \geq a) \leq \frac{E|X|}{a}$.
 - ◆ Chebyshev: $P(|X - \mu| \leq a) \leq \frac{\sigma^2}{a^2}$.
 - ◆ Chernoff: $P(X \geq a) \leq \frac{E(e^{tX})}{e^{ta}}$.
- ◆ Strong Law of Large #s (SLLN): $\bar{X}_n \rightarrow \mu$ pointwise, with prob 1.
- ◆ Weak Law of Large #s (WLLN): For all $\varepsilon > 0$, $P(|\bar{X}_n - \mu| > \varepsilon) \rightarrow 0$ as $n \rightarrow \infty$.



§10.3 - Central Limit Thm (CLT)



As in the previous section, let X_1, X_2, X_3, \dots be iid w/mean μ and variance σ^2 .

SLLN says that as $n \rightarrow \infty$, \bar{X}_n converges to the constant μ (w/prob 1).

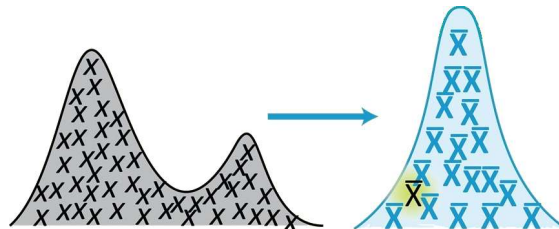
But what is its distr along the way to becoming a constant?

CLT states that for large n , the distr of \bar{X}_n (after standardization) approaches a standard Normal distr.

By standardization, we mean that we subtract μ (the mean of \bar{X}_n), and divide by $\frac{\sigma}{\sqrt{n}}$ (the SD of \bar{X}_n): $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$.

[see animation in class]

<https://for.instructure.com/resources/de99256fc78d4bbb16aec9b38acbf17ehared>



Thm (Central Limit Thm): As $n \rightarrow \infty$, we have $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \rightarrow \mathcal{N}(0, 1)$ in distr.

[Proof: in textbook]

In other words, this means the CDF of the LHS converges to Φ , the CDF of the standard Normal distr.

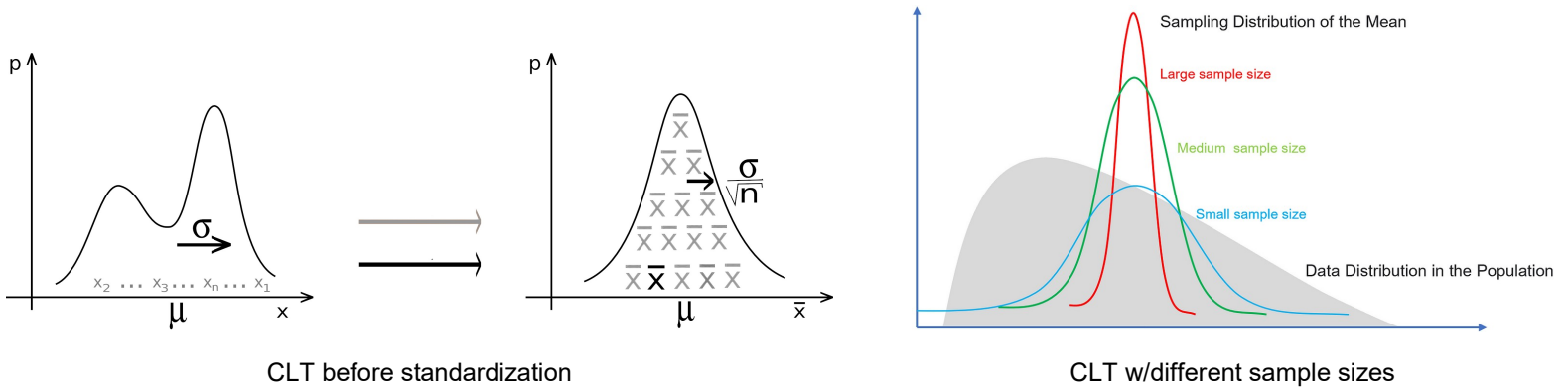
CLT suggests an approximation for the distr of X_n when n is large but finite:

Thm (CLT, Approximation Form): For large n , the distr of \bar{X}_n is approximately $\mathcal{N}(\mu, \frac{\sigma^2}{n})$.

Proof. Change the arrow in the CLT to \sim , an approximate distr sign: $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$

Then by location-scale transformation, $\bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. ■

We already saw from properties of expectation and variance that \bar{X}_n has mean μ and variance $\frac{\sigma^2}{n}$; CLT, however, tells us that \bar{X}_n is approximately Normal.



Ex (Running Proportion of Heads, Revisited):



Let X_1, X_2, \dots be iid $Bern(\frac{1}{2})$. LLN says that $\bar{X}_n \rightarrow \frac{1}{2}$ as $n \rightarrow \infty$.

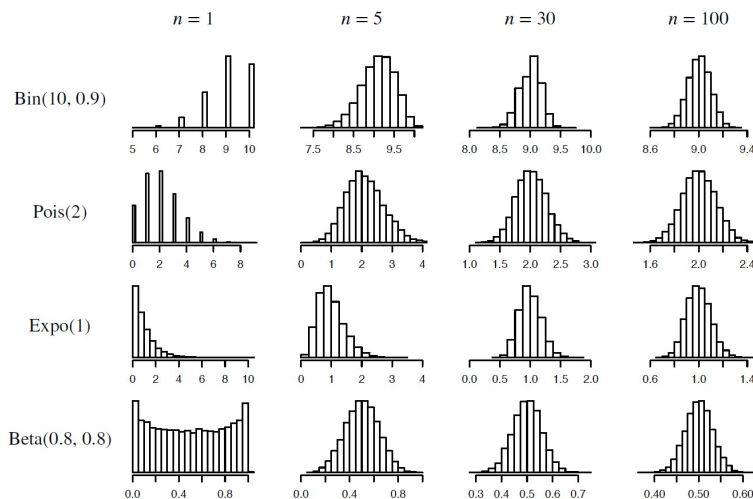
Now, using CLT, we can say more:

$$E(\bar{X}_n) = \frac{1}{2} \text{ and } Var(\bar{X}_n) = \frac{\sigma^2}{n} = \frac{p(1-p)}{n} = \frac{1}{4n}. \text{ And for large } n, \bar{X}_n \sim \mathcal{N}(\frac{1}{2}, \frac{1}{4n}).$$

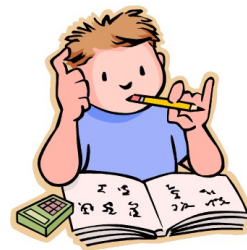
This lets us quantify a typical deviation from the mean is typical for a given n .

For example, when $n = 100$, $SD(\bar{X}_n) = \sqrt{\frac{1}{4 \cdot 100}} = \frac{1}{20} = 0.05$.

So, by the 68/95/99.7% rule there's an approx 95% chance that \bar{X}_n will be in $[0.4, 0.6]$ (2SDs).



Ex (Sampling Probability): As you do your homework, the time you spend on a problem follows an exponential distr, with a mean of 10 mins. You complete 36 problems, what's the prob that the average time you spend on those problems is between 9 and 11 mins?



We want: $P(9 < \bar{X}_{36} < 11)$.

Recall $\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) = \mathcal{N}\left(10, \frac{\sigma^2}{36}\right)$. Note our rate is $\lambda = \frac{1}{10}$ (per unit time, min).

Recall that $\sigma = \mu = 10$ for exponential distr. Why?

Comparing standard exponential X w/ $\lambda_X = 1$ and $Var(X) = 1$ against exponential Y w/ $\lambda_Y = \frac{1}{10}$: we have $Y = \frac{X}{\lambda_Y}$.

So, $\sigma^2 = Var(Y) = Var\left(\frac{X}{\lambda}\right) = \frac{1}{\lambda^2} Var(X) = \frac{1}{\lambda^2}$, so $\sigma = \frac{1}{\lambda} = 10 = \mu$.

So now we can calculate: $P(9 < \bar{X}_{36} < 11)$

$$= P\left(\frac{9-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X}_{16}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{11-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\frac{9-10}{\frac{10}{6}} < Z < \frac{11-10}{\frac{10}{6}}\right)$$

$$= P\left(-\frac{3}{5} < Z < \frac{3}{5}\right) = P\left(Z < \frac{3}{5}\right) - P\left(Z < -\frac{3}{5}\right)$$

$$= \Phi\left(\frac{3}{5}\right) - \Phi\left(-\frac{3}{5}\right). \quad \square$$

Activity 19

What did we learn?

- ◆ CLT: As $n \rightarrow \infty$, $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \rightarrow \mathcal{N}(0, 1)$
- ◆ CLT, Approximation Form: For large n , distr of \bar{X}_n is approximately $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$

