COMPSCI 240: Reasoning Under Uncertainty

Andrew Lan and Nic Herndon

University of Massachusetts at Amherst

Spring 2019

Lecture 20: Central limit theorem & The strong law of large numbers

Markov and Chebyshev Bounds

- **Markov Bound**
  - Informally: If a nonnegative RV has a small mean, then the probability that this RV takes a large value must also be small.
  - Formally: For a non-negative random variable X,

$$P(X \geq a) \leq \frac{E(X)}{a}$$

- **Chebyshev Bound**
  - Informally: If a RV has small variance, then the probability that it takes a value far from its mean is also small. Note that the Chebyshev inequality does not require the random variable to be nonnegative.
  - Formally: For a random variable X,

$$P(|X - E(X)| \geq c) \leq \frac{Var(X)}{c^2}$$

The Weak Law of Large Numbers

- Informally: If n is large, the bulk of the distribution of the sample mean ($\overline{X}_n$) of a sequence of i.i.d. with mean μ and variance σ^2 will converge to (be concentrated around) μ.
- Formally: Let $X_1, X_2, \cdots$ be a sequence of i.i.d. (either discrete or continuous) random variable with mean μ. For every $\epsilon > 0$, we have

$$P\left(|\overline{X}_n - \mu| \geq \epsilon\right) \to 0 \text{ as } n \to \infty.$$

Convergence in probability

- Let $Y_1, Y2, \ldots$ be a sequence of random variables (not necessarily independent), and let a be a real number.
- We say that the sequence Y_n **converges to** a **in probability**, if for every $\epsilon > 0$, we have

$$\lim_{n \to \infty} P(|Y_n - a| \geq \epsilon) = 0$$

- Put it another way: $\forall \epsilon, \delta > 0, \exists n_0$ such that $\forall n \geq n_0$

$$P(|Y_n - a| \geq \epsilon) \leq \delta$$

Our measurement is accurate, with this much confidence.

The Strong Law of Large Numbers

- Let $X_1, X_2, \cdots$ be a sequence of i.i.d. (either discrete or continuous) random variable with mean μ and variance σ^2.
- Then, the sequence of sample mean $\overline{X}_n$ converges to μ as $n \to \infty$, with probability 1:

$$P\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1.$$

- Its sample mean $\overline{X}_n$, which is a RV, will **converge** to the true mean μ, which is a constant, with a probability 1 when we have an infinitely large sample size.
  - More specifically, an event of $\overline{X}_n = \mu$ has a probability of 1.
- **Example**: Let $X_i \sim \text{Bern}(p)$, then

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = p\right) = 1.$$

The Central Limit Theorem

- The LLN states that $\overline{X}_n$ converges to μ when n is large.
  - The distribution of the sample mean $\overline{X}_n$ is concentrated around μ.
- But what does the distribution of $\overline{X}_n$ look like?
- **The Central Limit Theorem** can define this.

The Central Limit Theorem

- Let us define a variable by normalizing $\overline{X}_n$ with its mean and standard deviation
  - In the same manner as we normalized a Normal RV to derive the Standard Normal RV.

$$Z_n = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

  or equivalently

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

- Then, the PDF of Z_n converges to the standard normal PDF as $n \to \infty$

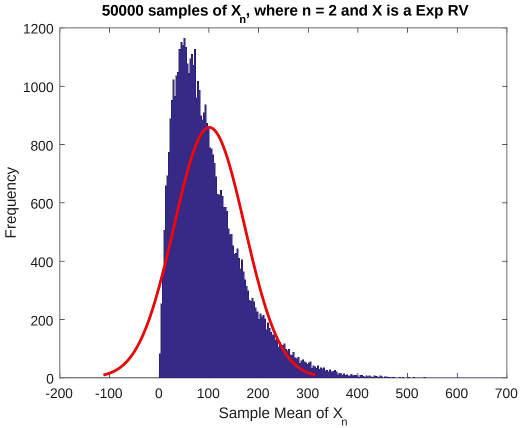
$$Z_n \sim N(0, 1) \text{ as } n \to \infty$$

The Central Limit Theorem

- The CLT is surprisingly general and extremely powerful.
- It states that X_i can have any forms of (discrete, continuous, or a mixture) probability distribution, but its sample mean converges to a Standard Normal distribution as n becomes large.
- Conceptually, this is important as it indicates that the sum of a large number of i.i.d RV is approximately normal.
- Practically, this is important as it eliminates the need for detailed probabilistic models as long as we have a large sample size. We can still approximate its sample mean using the Standard Normal distribution as long as we know μ and σ.

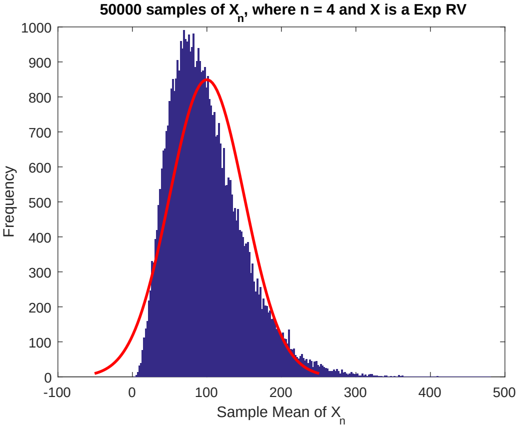
The Central Limit Theorem

- Let us run a simulation to see if this work!
- Consider a continuous exponential RV whose $\lambda = 0.01$
- Sampling distribution of $\overline{X}_n$ when $n = 2$



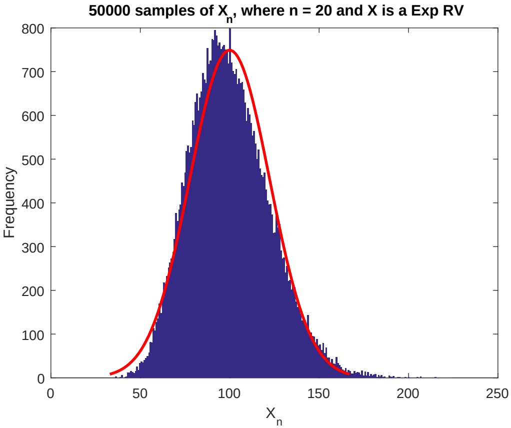
The Central Limit Theorem

- Let us run a simulation to see if this work!
- Consider a continuous exponential RV whose $\lambda = 0.01$
- Sampling distribution of $\overline{X}_n$ when $n = 4$



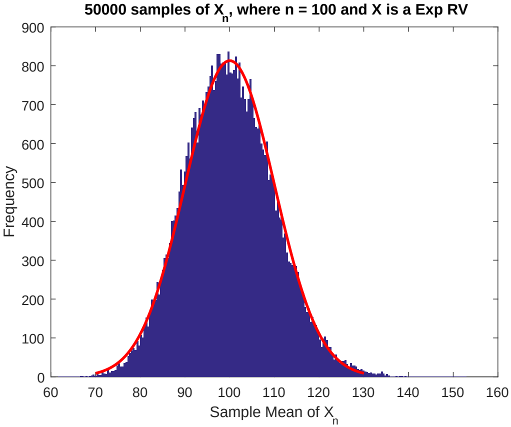
The Central Limit Theorem

- Let us run a simulation to see if this work!
- Consider a continuous exponential RV whose $\lambda = 0.01$
- Sampling distribution of $\overline{X}_n$ when $n = 20$



The Central Limit Theorem

- Let us run a simulation to see if this work!
- Consider a continuous exponential RV whose $\lambda = 0.01$
- Sampling distribution of $\overline{X}_n$ when $n = 100$



Example

- **Question**: Suppose salaries at a very large company have a mean of \$62,000 and a standard deviation of \$32,000.
- If a single employee is randomly selected, what is the probability that his/her salary exceeds \$66,000?
- **Solution**: We cannot solve this problem since we do not have the true distribution function of the salaries.

Example

- **Question**: Suppose salaries at a very large company have a mean of \$62,000 and a standard deviation of \$32,000.
- If 100 employees are randomly selected, what is the probability that their average salary exceeds \$66,000?
- **Solution**:
  - We define a new random variable

$$Z = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

    Then,

$$\begin{aligned} P(\overline{X}_n > 66000) &= P\left(Z > \frac{66000 - 62000}{\frac{32000}{\sqrt{100}}}\right) \\ &= P(Z > 1.25) \\ &= 1 - \Phi(1.25) \end{aligned}$$