Outline

1. Define convergent series
2. Definite the Law of Large Numbers
3. Define the Central Limit Theorem
4. Create Wald confidence intervals using the CLT
Numerical limits

- Imagine a sequence
  \[ a_1 = .9, \]
  \[ a_2 = .99, \]
  \[ a_3 = .999, \ldots \]

- Clearly this sequence converges to 1

- Definition of a limit: For any fixed distance we can find a point in the sequence so that the sequence is closer to the limit than that distance from that point on

- \[ |a_n - 1| = 10^{-n} \]
Limits of random variables

- The problem is harder for random variables
- Consider $\bar{X}_n$ the sample average of the first $n$ of a collection of iid observations
  
  Example $\bar{X}_n$ could be the average of the result of $n$ coin flips (i.e. the sample proportion of heads)

- We say that $\bar{X}_n$ converges in probability to a limit if for any fixed distance the probability of $\bar{X}_n$ being closer (further away) than that distance from the limit converges to one (zero)

- $P(|\bar{X}_n - \text{limit}| < \epsilon) \to 1$
The Law of Large Numbers

• Establishing that a random sequence converges to a limit is hard

• Fortunately, we have a theorem that does all the work for us, called the Law of Large Numbers

• The law of large numbers states that if $X_1, \ldots, X_n$ are iid from a population with mean $\mu$ and variance $\sigma^2$ then $\bar{X}_n$ converges in probability to $\mu$

• (There are many variations on the LLN; we are using a particularly lazy one)
Proof using Chebyshev’s inequality

- Recall Chebyshev’s inequality states that the probability that a random variable variable is more than $k$ standard deviations from its mean is less than $1/k^2$

- Therefore for the sample mean

$$P \left\{ |\bar{X}_n - \mu| \geq k \cdot sd(\bar{X}_n) \right\} \leq 1/k^2$$

- Pick a distance $\epsilon$ and let $k = \epsilon/sd(\bar{X}_n)$

$$P(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{sd(\bar{X}_n)^2}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$
Useful facts

- Functions of convergent random sequences converge to the function evaluated at the limit.
- This includes sums, products, differences, ...
- Example \((\bar{X}_n)^2\) converges to \(\mu^2\).
- Notice that this is different than \((\sum X_i^2)/n\) which converges to \(E[X_i^2] = \sigma^2 + \mu^2\).
- We can use this to prove that the sample variance converges to \(\sigma^2\).
Continued

\[
\sum (X_i - \bar{X}_n)^2 / (n - 1) = \frac{\sum X_i^2}{n - 1} - \frac{n(\bar{X}_n)^2}{n - 1}
\]

\[
= \frac{n}{n - 1} \times \frac{\sum X_i^2}{n} - \frac{n}{n - 1} \times (\bar{X}_n)^2
\]

\[
p \rightarrow 1 \times (\sigma^2 + \mu^2) - 1 \times \mu^2
\]

\[
= \sigma^2
\]

Hence we also know that the sample standard deviation converges to \(\sigma\)
Discussion

• An estimator is **consistent** if it converges to what you want to estimate

• The LLN basically states that the sample mean is consistent

• We just showed that the sample variance and the sample standard deviation are consistent as well

• Recall also that the sample mean and the sample variance are unbiased as well

• (The sample standard deviation is not unbiased, by the way)
The Central Limit Theorem

- The Central Limit Theorem (CLT) is one of the most important theorems in statistics.
- For our purposes, the CLT states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases.
- The CLT applies in an endless variety of settings.
The CLT

• Let $X_1, \ldots, X_n$ be a collection of iid random variables with mean $\mu$ and variance $\sigma^2$

• Let $\bar{X}_n$ be their sample average

• Then

$$P \left( \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \leq z \right) \rightarrow \Phi(z)$$

• Notice the form of the normalized quantity

$$\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} = \frac{\text{Estimate} - \text{Mean of estimate}}{\text{Std. Err. of estimate}}.$$
Example

- Simulate a standard normal random variable by rolling $n$ (six sided)
- Let $X_i$ be the outcome for die $i$
- Then note that $\mu = E[X_i] = 3.5$
- $\text{Var}(X_i) = 17.5/6$
- $\text{SE} \sqrt{17.5/6n} = 4.1833/\sqrt{6n}$
- Standardized mean
  \[ \frac{\bar{X}_n - 3.5}{4.1833/\sqrt{6n}} \]
- Notice that when $n = 6$ this formula simplifies to
  \[ \frac{\sum_{i=1}^{6} X_i - 21}{4.1833} \]
Coin CLT

- Let $X_i$ be the 0 or 1 result of the $i^{th}$ flip of a possibly unfair coin
- The sample proportion, say $\hat{p}$, is the average of the coin flips
- $E[X_i] = p$ and $\text{Var}(X_i) = p(1 - p)$
- Standard error of the mean is $p(1 - p)/\sqrt{n}$
- Then

$$\frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$$

will be approximately normally distributed
CLT in practice

- In practice the CLT is mostly useful as an approximation

\[ P \left( \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \leq z \right) \approx \Phi(z). \]

- Recall 1.96 is a good approximation to the .975th quantile of the standard normal

- Consider

\[ .95 \approx P \left( -1.96 \leq \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \leq 1.96 \right) \]

\[ = P \left( \bar{X}_n + 1.96\sigma / \sqrt{n} \geq \mu \leq \bar{X}_n - 1.96\sigma / \sqrt{n} \right), \]
Confidence intervals

- Therefore, according to the CLT, the probability that the random interval
  \[ \bar{X}_n \pm z_{1-\alpha/2} \sigma / \sqrt{n} \]
  contains \( \mu \) is approximately 95%, where \( z_{1-\alpha/2} \) is the \( 1 - \alpha/2 \) quantile of the standard normal distribution

- This is called a 95% confidence interval for \( \mu \)

- Slutsky’s theorem, allows us to replace the unknown \( \sigma \) with \( s \)
Sample proportions

- In the event that each $X_i$ is 0 or 1 with common success probability $p$ then $\sigma = p(1 - p)$
- The interval takes the form
  \[ \hat{p} \pm z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}} \]
- Replacing $p$ by $\hat{p}$ in the standard error results in what is called a Wald confidence interval for $p$
- Also note that $p(1 - p) \leq 1/4$ for $0 \leq p \leq 1$
- Let $\alpha = .05$ so that $z_{1-\alpha/2} = 1.96 \approx 2$ then
  \[ 2 \sqrt{\frac{p(1-p)}{n}} \leq 2 \sqrt{\frac{1}{4n}} = \frac{1}{\sqrt{n}} \]
- Therefore $\hat{p} \pm \frac{1}{\sqrt{n}}$ is a quick CI estimate for $p$