5. Limit Theorems, Part I: Weak Law of Large Numbers

ECE 302 Spring 2012
Purdue University, School of ECE
Prof. Ilya Pollak
Some motivation

• Suppose $X_1, \ldots, X_n$ are i.i.d. (independent, identically distributed) with mean $\mu$ and variance $\sigma^2$
• The sample mean is $M_n = (X_1 + \ldots + X_n)/n$
• What happens as $n \to \infty$ ?
• Why is answering this question important?
  – Will allow us to interpret expectations and probabilities in terms of a long sequence of independent identical experiments.
  – Will facilitate approximate analysis of r.v.’s such as $M_n$ whose distribution is difficult to compute for large $n$. 

Ilya Pollak
Weak law of large numbers

• Suppose \( X_1, \ldots, X_n \) are i.i.d. (independent, identically distributed) with mean \( \mu \) and variance \( \sigma^2 \)

• The sample mean is \( M_n = (X_1 + \ldots + X_n)/n \)

\[
P(|M_n - \mu| \geq \varepsilon) \to 0 \text{ as } n \to \infty, \text{ for every } \varepsilon > 0
\]

• This says that the bulk of the distribution of \( M_n \) is concentrated near the actual mean \( \mu \).
Plan for proving WLLN

• Tools: Markov and Chebyshev inequalities.
• Meaning of “→”
Markov Inequality

If \( P(X < 0) = 0 \), then

\[
P(X \geq a) \leq \frac{E[X]}{a}, \quad \text{for all } a > 0.
\]

**Proof.** Fix \( a > 0 \), and let \( g_a(x) = \begin{cases} 0, & \text{if } 0 \leq x < a \\ a, & \text{if } x \geq a \end{cases} \)

Then \( g_a(x) \leq x \) for all \( x \geq 0 \), and so

\[
E[g_a(X)] \leq E[X]
\]

But \( E[g_a(X)] = 0 \cdot P(X < a) + a \cdot P(X \geq a) = a \cdot P(X \geq a) \)

Hence, \( a \cdot P(X \geq a) \leq E[X] \).
Chebyshev Inequality

If $E[X] = \mu$ and $\text{var}(X) = \sigma^2$, then

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}, \text{ for any } c > 0.$$  

Alternative form:  

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}, \text{ for any } k > 0.$$  

Proof. Since $(X - \mu)^2 \geq 0$ with probability 1, we can apply Markov inequality. Let $a = c^2$. Then

$$P((X - \mu)^2 \geq c^2) \leq \frac{E[(X - \mu)^2]}{c^2} = \frac{\sigma^2}{c^2}.$$  

But $P((X - \mu)^2 \geq c^2) = P(|X - \mu| \geq c)$.  

If $c = k\sigma$, then

$$P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2 \sigma^2} = \frac{1}{k^2}.$$  

Ilya Pollak
Review of deterministic limits

Suppose \(a_1, a_2, a_3, \ldots\) is a sequence of numbers.

A number \(a\) is the limit of \(a_n\) (or \(a_n\) converges to \(a\)) if, for every \(\delta > 0\), there exists \(n_0\) such that for all \(n \geq n_0\) we have \(|a_n - a| \leq \delta\).

Notation: \(a_n \to a\) as \(n \to \infty\), or \(\lim_{n \to \infty} a_n = a\).
Convergence in probability

Suppose \( Y_1, Y_2, Y_3, \ldots \) is a sequence of random variables. \( Y_n \) converges in probability to a number \( a \) if, for every \( \epsilon > 0 \),

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

---in other words, if, for every \( \delta > 0 \) and \( \epsilon > 0 \), there exists \( n_0 \) such that for all \( n \geq n_0 \),

\[
P(\{|Y_n - a| \geq \epsilon\}) \leq \delta.
\]

Interpretation: For any given levels of accuracy \( \epsilon \) and confidence \( \delta \), the random variable \( Y_n \) is likely to be approximately equal to \( a \), within these levels of accuracy and confidence, provided that \( n \) is large enough.
Ex. 5.8: Convergence in probability

\[ p_{Y_n}(y) \]

\[
\begin{array}{c}
\text{0} \quad \text{1} - \frac{1}{n} \\
\text{n}^{2} \quad \frac{1}{n}
\end{array}
\]

\[
P(|Y_n - 0| \geq \varepsilon) = \begin{cases} 
0, & \varepsilon > n^2 \\
\frac{1}{n}, & 0 < \varepsilon \leq n^2
\end{cases}
\leq \frac{1}{n} \rightarrow 0 \text{ as } n \rightarrow \infty, \text{ for any } \varepsilon > 0.
\]

Therefore, \( Y_n \) converges in probability to zero.

But note: \( E[Y_n] = n^2 \cdot \frac{1}{n} = n \rightarrow \infty \text{ as } n \rightarrow \infty. \)
Weak law of large numbers

• Suppose $X_1, \ldots, X_n$ are i.i.d. (independent, identically distributed) with mean $\mu$ and variance $\sigma^2$

• The sample mean $M_n = (X_1 + \ldots + X_n)/n$ converges to the actual mean $\mu$ in probability:

$$\Pr(|M_n - \mu| \geq \varepsilon) \to 0 \quad \text{as} \quad n \to \infty,$$

for every $\varepsilon > 0$
Weak law of large numbers: Proof

• $X_1, \ldots, X_n$ are i.i.d. with mean $\mu$ and variance $\sigma^2$
• The sample mean $M_n = (X_1 + \ldots + X_n)/n$ converges to the actual mean $\mu$ in probability.

Proof. Chebyshev inequality:

$$P\left(|M_n - E[M_n]| \geq \varepsilon\right) \leq \frac{\text{var}(M_n)}{\varepsilon^2}.$$ 

But $E[M_n] = \mu$ and $\text{var}(M_n) = \frac{\sigma^2}{n}$. Therefore,

$$P\left(|M_n - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \to 0 \quad \text{as} \quad n \to \infty, \text{ for every } \varepsilon > 0.$$
Ex. 5.4: Probabilities and Frequencies

• Suppose that during each repetition of an experiment, event $A$ occurs with probability $p = P(A)$.

• Consider $n$ independent repetitions.

• Let $M_n = (\text{number of occurrences of } A)/n$

Let $X_i = \begin{cases} 1, & \text{if } A \text{ occurs on the } i\text{-th experiment} \\ 0, & \text{if } A \text{ does not occur on the } i\text{-th experiment} \end{cases}$

Then $X_i$'s are independent Bernoulli random variables, with

$$p_{X_i}(x_i) = \begin{cases} p, & \text{if } x_i = 1 \\ 1 - p, & \text{if } x_i = 0 \end{cases}$$

$$M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

$$E[X_i] = p, \text{ hence } E[M_n] = p$$

WLLN: when $n$ is large, the empirical frequency $M_n$ of $A$ is very likely to be close to $p$, the probability of $A$. 
Ex. 5.5: Polling

Estimate President Obama's approval rating by asking \( n \) persons drawn at random from the voter population. Let

\[
X_i = \begin{cases} 
1, & \text{if the } i\text{-th person approves} \\
0, & \text{otherwise}
\end{cases}
\]

Model \( X_1, X_2, \ldots, X_n \) as independent Bernoulli r.v.'s with mean \( p \) and variance \( p(1-p) \), where \( p \) is the "true" approval rating.

We estimate \( p \) as the sample mean: 

\[
M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}
\]

\[
E[M_n] = p, \quad \text{var}(M_n) = \frac{p(1-p)}{n}
\]

By WLLN, \( M_n \to p \) in probability as \( n \to \infty \).

By Chebyshev inequality, \( \mathbb{P}(|M_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \)
Ex. 5.5: Polling

Estimate President Obama's approval rating by asking $n$ persons drawn at random from the voter population. Let

$$X_i = \begin{cases} 
1, & \text{if the } i\text{-th person approves} \\
0, & \text{otherwise}
\end{cases}$$

Model $X_1, X_2, \ldots, X_n$ as independent Bernoulli r.v.'s with mean $p$ and variance $p(1-p)$, where $p$ is the "true" approval rating.

We estimate $p$ as the sample mean: $M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$

$$E[M_n] = p, \quad \text{var}(M_n) = \frac{p(1-p)}{n}$$

By WLLN, $M_n \to p$ in probability as $n \to \infty$.

By Chebyshev inequality, $P\left(\left|M_n - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$

Suppose we want $P\left(\left|M_n - p\right| \geq 0.01\right) \leq 0.05$. I.e., we want to be 95% confident that we are within 0.01 of the actual approval rating. We can guarantee this if we take $n \geq \frac{1}{4\varepsilon^2 0.05} = \frac{1}{4 \cdot 0.01^2 \cdot 0.05} = 50,000$. 

Ilya Pollak
Ex. 5.5: Polling

Estimate President Obama's approval rating by asking \( n \) persons drawn at random from the voter population. Let

\[
X_i = \begin{cases} 
1, & \text{if the } i-\text{th person approves} \\
0, & \text{otherwise}
\end{cases}
\]

Estimate \( p \), the "true" approval rating, as the sample mean: 

\[
M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}
\]

\[
E[M_n] = p, \quad \text{var}(M_n) = \frac{p(1-p)}{n}
\]

By WLLN, \( M_n \to p \) in probability as \( n \to \infty \).

By Chebyshev inequality, 

\[
P(|M_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}
\]

Suppose we want 

\[
P(|M_n - p| \geq 0.01) \leq 0.05.
\]

I.e., we want to be 95% confident that we are within 0.01 of the actual approval rating. We can guarantee this if we take 

\[
n \geq \frac{1}{4\varepsilon^20.05} = \frac{1}{4 \cdot 0.01^2 \cdot 0.05} = 50,000.
\]

Note: This is very conservative because the Chebyshev inequality is loose!