STAT 346/446 Lecture 2

Convergence concepts

CB Sections 5.5.1 - 5.5.3
DS Sections 6.1 - 6.2

Convergence Concepts

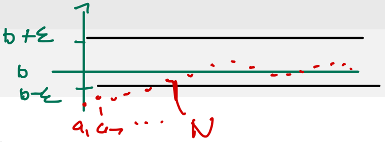
- Behavior of sample statistics as sample size approaches infinity
    - “somewhat fanciful idea”
    - Mostly used to find approximations to use for finite sample sizes
- We are interested in sequences of random variables

$$X_1, X_2, X_3, \ldots$$

    - Do the distributions of X_n converge to a limiting distribution?
- In particular we are very interested in the sequence

$$\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots \qquad \text{where} \qquad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Limits - review



- Limit of a sequence of real numbers, $a_1, a_2, a_3, \ldots$

$$\lim_{n\to\infty} a_n = b \qquad \text{or:} \qquad a_n \to b \text{ as } n \to \infty$$

The sequence can get ε close for N large enough

  - Meaning: For any real number $\varepsilon > 0$, there exists a natural number N such that for all $n > N$ we have $|a_n - b| < \varepsilon$.

- Limit of a function $f : \mathbb{R} \to \mathbb{R}$

$$\lim_{x\to c} f(x) = b \qquad \text{or:} \qquad f(x) \to b \text{ as } x \to c$$

  - Meaning: For any real number $\varepsilon > 0$, there exists a real number $\delta > 0$ such that for all x such that $|x - c| < \delta$ we have $\underbrace{|f(x) - b|}_{\varepsilon \text{ close}} < \varepsilon$.

Can get ε close for δ small enough (x close enough to c)

Limits - for random variables

- What does a limit for a sequence of random variables mean?

$$\lim_{n\to\infty} X_n = X$$

  - Meaning: ?

- $|X_n - X|$ is a random variable
- Or do we just care about the limits of the distributions of X_n i.e. the cdfs $F_n(x)$?
- Many ways of defining convergence for random variables

Different Convergence Concepts

- **Converges in probability** Section 5.5.1
- **Converges almost surely** Section 5.5.2
- **Converges in distribution** Section 5.5.3
- Weak and strong law of large numbers
- Central Limit Theorem!
- Slutskys's Theorem
- Delta Method Section 5.5.4

next lecture

Convergence Concept: In probability

Def: Convergence in probability

A sequence of random variables $X_1, X_2, X_3, \ldots$ **converges in probability** to a random variable X if for any real number $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

$$\Rightarrow \lim_{n\to\infty} [1 - P(|X_n - X| < \varepsilon)] = 0$$

or equivalently:

$$\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

- Notation for convergence in probability:

$$X_n \xrightarrow{p} X \quad \text{as } n \to \infty$$

Convergence in probability

sequence of
$p_1, p_2, p_3, \ldots$

- Things to note about the definition

$$\lim_{n\to\infty} \overbrace{P\left(|X_n - X| \geq \varepsilon\right)}^{p_n} = 0$$

  - $P\left(|X_n - X| \geq \varepsilon\right)$ is a sequence of numbers
    - So lim is defined
  - $P\left(|X_n - X| \geq \varepsilon\right)$ involves two random variables $\Rightarrow$ need the joint distribution of X_n and X to evaluate

- We are often interested in the special case when X is a constant, i.e.

$$X_n \xrightarrow{p} a \quad \text{as } n \to \infty$$

i.e. $\lim_{n\to\infty} P(|X_n - a| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0$

Weak law of large numbers

WLLN

Let $X_1, X_2, X_3, \ldots$ be a sequence of iid. random variables with $\mathrm{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2 < \infty$. Define

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Then:

$$\overline{X}_n \xrightarrow{P} \mu \quad \text{as } n \to \infty$$

Proof...

on the whiteboard

$\overline{X}_n$ is a consistent estimator of μ (any distr.!)

Consistent estimators

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

- Since $\overline{X}_n \xrightarrow{p} \mu$ we say that $\overline{X}_n$ is a **consistent estimator** of μ
- Can also show that S_n^2 is a consistent estimator of σ^2 for most distributions (see textbook)

$$P(|S_n^2 - \sigma^2| \geq \varepsilon) = P((S_n^2 - \sigma^2)^2 \geq \varepsilon^2)$$

$$\leq \frac{E((S_n^2 - \sigma^2)^2)}{\varepsilon^2} = \frac{Var(S_n^2)}{\varepsilon^2}$$

All we need is $Var(S_n^2) \to 0$ as $n \to \infty$

In fact $Var(S_n^2) = \frac{1}{n}\left(\theta_4 - \frac{n-3}{n-1}\theta_2^2\right)$

where $\theta_4 = E((X-\mu)^4)$ and $\theta_2 = E((X-\mu)^2)$

$\Rightarrow$ If $\theta_4 < \infty$ then $Var(S_n^2) \to 0$

More on convergence in probability

Theorem

If $X_n \xrightarrow{p} X$ and $h(x)$ is a continuous function then

$$h(X_n) \xrightarrow{p} h(X)$$

Example:

Since $\bar{X}_n \xrightarrow{P} \mu$ then $\bar{X}_n^2 \xrightarrow{P} \mu^2$

Since $S_n^2 \xrightarrow{P} \sigma^2$ then $S_n \xrightarrow{P} \sigma$

$S_n = \sqrt{S_n^2}$

Convergence Concept: Almost Surely

Def: Convergence almost surely

A sequence of random variables $X_1, X_2, X_3, \ldots$ **converges almost surely** to a random variable X if for any real number $\varepsilon > 0$,

$$P\left(\lim_{n\to\infty} |X_n - X| < \varepsilon\right) = 1$$

- Compare to conv. in prob.

$$\lim_{n\to\infty} P\left(|X_n - X| < \varepsilon\right) = 1$$

- Notation for convergence almost surely:

$$X_n \xrightarrow{\text{as}} X$$

Convergence almost surely

$P(\lim_{n\to\infty} | X_n(s) - X(s) | < \varepsilon)$

- Recall: A random variable X_n is a function from sample space to $\mathbb{R}$

$$P\left(\lim_{n\to\infty} |X_n - X| < \varepsilon\right) = 1$$

  - Meaning: $X_1(s), X_2(s), X_3(s), \ldots$ converges to $X(s)$

$$X_n(s) \to X(s)$$

  for all $s \in S$, except perhaps for outcomes s in a probability zero region

- Stronger convergence than in probability, i.e.

  but trickier to work with.

$$\text{If} \quad X_n \xrightarrow{\text{as}} X \quad \text{then} \quad X_n \xrightarrow{P} X$$

- Again, we are often interested in the special case when X is a constant, i.e.

$$X_n \xrightarrow{\text{as}} a$$

Strong law of large numbers

SLLN

Let $X_1, X_2, \ldots, X_n$ be a sequence of iid. random variables with $\mathrm{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2 < \infty$. Define

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Then

$$\overline{X}_n \xrightarrow{\text{as}} \mu$$

- Both WLLN and SLLN can actually be proven without the $\mathrm{Var}(X_i) < \infty$ assumption.

Laws of large numbers

Example

- Suppose $X_1, X_2, X_3, \ldots$ are i.i.d. $\mathrm{Bernoulli}(p)$ $\quad E(X_i) = p$

- Let

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad = \text{ sample proportion of success}$$

  i.e. $\overline{X}_n$ is the proportion of successes in n trials

- Weak and Strong LLN:

$$\overline{X}_n \xrightarrow{p} p$$
$$\overline{X}_n \xrightarrow{\text{as}} p$$

Convergence Concept: In distribution

Def: Convergence in distribution

A sequence of random variables $X_1, X_2, X_3, \ldots$ **converges in distribution** to a random variable X if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

for all x where $F_X(x)$ is continuous.

- Implied notation: $F_{X_n}(x)$ is the cdf of X_n, $F_X(x)$ is the cdf of X
- Notation for convergence in distribution:

$$X_n \xrightarrow{d} X$$

Convergence in distribution

- Convergence of functions (cdfs)

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

- If pdf/pmf exist and $\lim_{n\to\infty} f_{X_n}(x) = f_X(x)$ for all $x \in \mathbb{R}$ then $X_n \xrightarrow{d} X$

- Saw in STAT 345/445: If mgfs exist and

$$\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$$

  for all t in a neighborhood of 0, then $X_n \xrightarrow{d} X$

- mgfs (and pdfs) are sometimes easier to work with

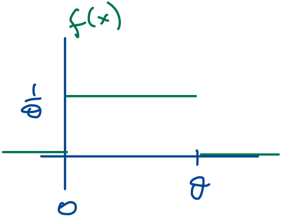
e.g. $F(x)$ for $N(\mu, \sigma^2)$ is not available in closed form but $M(t) = e^{\mu t + t^2\sigma^2/2}$

Example

- Let $X_1, X_2, \ldots$ be iid. $\mathrm{Uniform}(0, \theta)$ and let

$$X_{(n)} = \max_{1 \leq i \leq n} X_n$$

show that $X_{(n)} \xrightarrow{d} \theta$ and $X_{(n)} \xrightarrow{p} \theta$



On the whiteboard...

Weaker and stronger convergence concepts

Theorem: Conv almost surely $\Rightarrow$ Conv. in probability

If a sequence of random variables $X_1, X_2, X_3, \ldots$ converges almost surely to a random variable X the sequence converges also in probability to X.

- In compact notation:

$$\text{If} \quad X_n \xrightarrow{\text{as}} X \quad \text{then} \quad X_n \xrightarrow{p} X$$

Weaker and stronger convergence concepts

Theorem: Conv. in prob. $\Rightarrow$ Conv. in distribution

If a sequence of random variables $X_1, X_2, X_3, \ldots$ converges in probability to a random variable X the sequence converges also in distribution to X.

- In compact notation:

$$\text{If} \quad X_n \xrightarrow{p} X \quad \text{then} \quad X_n \xrightarrow{d} X$$

- We also get $\Leftarrow$ if X is a constant, i.e.

$$\text{If} \quad X_n \xrightarrow{d} a \quad \text{then} \quad X_n \xrightarrow{p} a$$

next slide ...

When "$\xrightarrow{d}$" implies "$\xrightarrow{p}$"

- Let $X_1, X_2, \ldots$ be a sequence of random variables and let a be a constant. Show that

$$X_n \xrightarrow{d} a \quad \Rightarrow \quad X_n \xrightarrow{p} a$$

on the board...