Section 2: Convergence Concepts

September 2015

1 Convergence Concepts

1.1 Converge in probability

1.1.1 Convergence of real numbers

- Recall that we say a sequence of real numbers $\{a_n : n \geq 1\}$ converges to a if $\forall \epsilon > 0$, there exists $N \geq 1$ such that
$$|a_n - a| \leq \epsilon$$
for every $n \geq N$.
- Similarly, we might want to define convergence for a sequence of random variables $\{X_n, n \in \mathcal{N}\}$ by looking at $|X_n - X|$, for some random variable X.
- Notice that $|X_n - X| \geq \epsilon$ is an random inequality.
- To define its convergence, we need to consider $P(|X_n - X| \geq \epsilon)$

1.1.2 Convergence in probability

- Let $\{X_n, n \in \mathcal{N}\}$ be a sequence of random variables, and let X be some random variable. We say that $\{X_n, n \in \mathcal{N}\}$ *converges in probability* to X if $\forall \epsilon > 0$
$$P(|X_n - X| \geq \epsilon) \to 0$$
- We denote it as $X_n \stackrel{P}{\to} X$.
- $X_n \stackrel{P}{\to} X$ does not imply $E(X_n) \to E(X)$!

1.2 Converge almost surely

- There is another type of convergence, which is also defined via the actual value of the random variables. and is called almost sure convergence.

- We say that $\{X_n\}$ *converges almost surely* to X if

$$P(\lim_{n\to\infty} X_n = X) = 1$$

Note that convergence in probability is defined by convergence of probabilities of a sequence of events. Is this the only choice?

1.3 Converge in distribution

- Let $\{X_n, n \in \mathcal{N}\}$ be a sequence of random variables, and let X be some random variable. Let $F^{(X_n)}$ and $F^{(X)}$ denote their cdfs. We say that $\{X_n\}$ *converges in distribution* to X if $F^{(X_n)}(x) \to F^{(X)}(x)$, for every x at which $F^{(X)}$ is continuous.
- We denote it as $X_n \stackrel{D}{\to} X$.
- Note that convergence in distribution is defined by convergence of cdfs, instead of the actual values taken by the random variables.

1.4 Relationships

- If $X_n \stackrel{a.e.}{\to} X$, then $X_n \stackrel{P}{\to} X$.
- If $X_n \stackrel{P}{\to} X$, then $X_n \stackrel{D}{\to} X$.
- Let $a \in R$ be a constant. Then $X_n \stackrel{P}{\to} a$, if and only if $X_n \stackrel{D}{\to} a$.

1.5 Convergence of continuous functions of random variables

One important property of convergence of random variables it that it is preserved in under continuous transformation

- If $X_n \stackrel{P}{\to} a$ for some constant a and $g : R \to R$ is continuous at a, then $g(X_n) \stackrel{P}{\to} g(a)$.
- If $X_n \stackrel{P}{\to} X$ and $g : R \to R$ is continuous, then $g(X_n) \stackrel{P}{\to} g(X)$.
- If $X_n \stackrel{D}{\to} X$ and $g : R \to R$ is continuous, then $g(X_n) \stackrel{D}{\to} g(X)$.

1.6 Slutsky's Theorem

If $X_n \stackrel{D}{\to} X$ and $Y_n \stackrel{P}{\to} a$, where $a \in R$ is a constant, then $X_n + Y_n \stackrel{D}{\to} X + a$ and $X_n Y_n \stackrel{D}{\to} aX$.

Note that to apply Slutsky's Theorem, the sequence of random variables Y_n must be converging to a constant.

2 Weak Law of Large Numbers and Central Limit Theorem

2.1 Weak Law of Large Numbers

2.1.1 Weak Law of Large Numbers

- Let $\{X_n, n \in \mathcal{N}\}$ be a sequence of *iid* random variables with $E(|X_i|) < \infty$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, then

$$\bar{X}_n \stackrel{P}{\rightarrow} E(X_i)$$

- WLLN formalizes the intuition that the expectation of a random variable may be interpreted as its long-run average.
- The theorem assumes nothing about the variance of X_i. In fact, it holds for the case where $Var(X_i) = \infty$.
- In fact, there also exits a strong law of large number which says that $\bar{X}_n \stackrel{a.e.}{\rightarrow} E(X_i)$.
- In more sophisticated versions of these theorems, the *iid* assumption can be relax much more for the for the weak law than for the strong law.
- However, we do need the assumption that $E(|X_i|) < \infty$.

2.2 Central Limit Theorem

2.2.1 Central Limit Theorem

- Let $\{X_n, n \in \mathcal{N}\}$ be a sequence of *iid* random variables with $Var(X_i) < \infty$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then

$$\sqrt{n}(\bar{X}_n - \mu) \stackrel{D}{\rightarrow} N(0, \sigma^2)$$

  .

  where $\mu = E(X_i)$ and $\sigma^2 = Var(X_i)$

- Informally, CLT states that for large n, $\bar{X}_n$ is approximated normal with mean μ and variance σ^2/n.
- The WLLN stated above is implied by the CLT. However, In more sophisticated versions of these theorems, the *iid* assumption can be relax much more for the for the WLLN than for the CLT.

3 Delta Method

3.1 Asymptotic of continuous functions of random variables

- Let $\{Y_n, n \geq 1\}$ be a sequence of random variables such that $\sqrt{n}(Y_n - a) \stackrel{D}{\to} Z$ for some random variable Z and some constant $a \in R$. Let $g : R \to R$ be a function. What can we say about the asymptotic behavior of $g(Y_n)$?

- We already know that $g(Y_n) \stackrel{D}{\to} g(a)$. Can we do better?

- If we assume g is differentiable at a, then

$$\sqrt{n}(g(Y_n) - g(a)) \approx g'(a)\sqrt{n}(Y_n - a) \stackrel{D}{\to} g'(a)Z$$

- This is basic idea behind *delta method*.

3.2 Delta Method

- Let $\{Y_n, n \geq 1\}$ be a sequence of random variables such that $\sqrt{n}(Y_n - a) \stackrel{D}{\to} Z$ for some random variable Z and some constant $a \in R$. Let $g : R \to R$ be a continuously differentiable at a. Then

$$\sqrt{n}(g(Y_n) - g(a)) \stackrel{D}{\to} g'(a)Z$$