Useful Probability Theorems

Shiu-Tang Li

Finished: March 23, 2013
Last updated: November 2, 2013

1 Convergence in distribution

Theorem 1.1. TFAE:

(i) \( \mu_n \Rightarrow \mu \), \( \mu_n, \mu \) are probability measures.
(ii) \( F_n(x) \to F(x) \) on each continuity point of \( F \).
(iii) \( E[f(X_n)] \to E[f(X)] \) for all bounded continuous function \( f \).
(iv) \( \phi_n(x) \to \phi(x) \) uniformly on every finite interval. [Breiman]
(v) \( \phi_n(x) \to \tilde{\phi}(x) \), \( \tilde{\phi} \) is continuous at \( 0 \).

Theorem 1.2. Let \( X_n \to c \) in dist, where \( c \) is a constant. Then \( X_n \to c \) in pr.

Theorem 1.3. Let \( X_n \to X \) in dist, \( Y_n \to 0 \) in dist, \( W_n \to a \) in dist, and \( Z_n \to b \) in dist, \( a, b \) are constants. Then

(1) \( aX_n + b \to aX + b \) in dist. (Consider continuity intervals)
(2) \( X_n + Y_n \to X \) in dist.
(3) \( X_n Y_n \to 0 \) in dist.
(4) \( W_n X_n + Z_n = aX_n + b + (W_n - a)X_n + (Z_n - b) \to aX + b \) in dist.

Remark. We only need to memorize (4). (Chung)

Theorem 1.4. If \( X_n \to X \) in dist, and \( f \in C(\mathbb{R}) \), then \( f(X_n) \to f(X) \) in dist. (Chung; consider bounded continuous test functions.)

Theorem 1.5. (Lindeberg-Feller; a generalization of CLT.) For each \( n \), let \( X_{n,m}, 1 \leq m \leq n \), be independent random variables with \( EX_{n,m} = \)
0. Suppose

(1) \[ \sum_{m=1}^{n} E[X_{n,m}^2] \to \sigma^2 > 0 \text{ as } n \to \infty \]
(2) For all \( \epsilon > 0 \), \( \lim_{n \to \infty} \sum_{m=1}^{n} E[X_{n,m}^2; |X_{n,m}| > \epsilon] = 0 \).

Then \( S_n = X_{n,1} + \cdots + X_{n,n} \) converges to \( N(0, \sigma^2) \) in distribution. (Durrett)

2 \quad \text{Convergence in probability}

Theorem 2.1. (Chebyshev's inequality.) If \( \phi \) is a strictly positive and increasing function on \((0, \infty)\), \( \phi(u) = \phi(-u) \), and \( X \) is an r.v. s.t. \( E[\phi(X)] < \infty \), then for each \( u \geq 0 \), we have \( P(|X| \geq u) \leq \frac{E[\phi(X)]}{\phi(u)}. \) (Chung)

Theorem 2.2. (Useful lower bound) Let \( X \geq 0 \), \( E[X^2] < \infty \), and \( 0 \leq a < E[X] \). We have \( P(X > a) \geq \frac{(EX - a)^2}{E[X^2]} \). (Durrett)

Theorem 2.3. Let \( X_n \to X \) in pr, \( Y_n \to Y \) in pr, \( f \in C(\mathbb{R}) \). Then \( f(X_n) \to f(X) \) in pr, \( X_n \pm Y_n \to X \pm Y \) in pr, and \( X_n Y_n \to XY \) in pr.

Theorem 2.4. \( X_n \to X \) pr. if and only if for every subsequence \( \{X_{n_j}\} \) of \( \{X_n\} \), we may pick a further subsequence \( \{X_{n_{j_k}}\} \) so that \( X_{n_{j_k}} \to X \) a.s. (Durrett; theorem 1.6.2)

3 \quad \text{Convergence a.s.}

Theorem 3.1. (Borel-Cantelli Lemma.)

Theorem 3.2. (Komolgorv's maximal inequality) Suppose \( S_n = X_1 + \cdots + X_n \), where the \( X_j \)'s are independent and in \( L^2(P) \). Then for all \( \lambda > 0 \) and \( n \geq 1 \), \( P(\max_{1 \leq k \leq n} |S_k - ES_k| > \lambda) \leq \frac{Var(S_n)}{\lambda^2}. \)

Theorem 3.3. (Kronecker's Lemma.) Let \( \{X_k\} \) be a sequence of real numbers, \( \{a_k\} \) a positive sequence that goes to infinity. Then \( \sum_{n=1}^{\infty} \frac{x_n}{a_n} \) converges \( \Rightarrow \frac{1}{a_n} \sum_{j=1}^{n} x_j \to 0 \). (Chung)

Theorem 3.4. Let \( \{X_n\} \) be a sequence of independent r.v.'s with \( E[X_n] = 0 \) for every \( n \), and \( 0 < a_n \uparrow \infty \). If \( \phi \in C(\mathbb{R}) \) is chosen s.t. \( \phi(x)/|x| \uparrow, \phi(x)/x^2 \downarrow \), as \( |x| \) increases, \( \phi \) is positive and even, and \( \sum_{n=1}^{\infty} \frac{E[\phi(X_n)]}{\phi(a_n)} \)
converges, then $\sum_{n=1}^{\infty} \frac{X_n}{a_n}$ converges a.s. (Chung; how to use this?)

**Theorem 3.5.** Let $\{X_n\}$ be i.i.d and $p > 0$. Then TFAE: (exercise 6.10 of Davar)

1. $E|X_n|^p < \infty$.
2. $X_n \in o(n^{1/p})$ a.s.
3. $\max_{1 \leq j \leq n} |X_j| \in o(n^{1/p})$ a.s.

**Theorem 3.6.** If $\{X_n\}$ are i.i.d and in $L^1(P)$, then

$$\lim_{n \to \infty} \frac{1}{\ln(n)} \sum_{j=1}^{\infty} \frac{X_j}{j} = EX_1 \text{ a.s.}$$

(exercise 6.29 of Davar)

**Theorem 3.7.** Let $\{X_n\}$ be i.i.d with $EX_n = 0$, $S_n := X_1 + \cdots + X_n$, and we assume that $E|X_n|^p < \infty$. Then:

1. When $p = 1$, $S_n \to 0$. (SLLN)
2. When $1 < p \leq 2$, $\frac{S_n}{n^{1/p}} \to 0$ a.s.
3. When $p = 2$, \( \frac{S_n}{n^{1/2}(\log(n))^{1/2+\epsilon}} \to 0 \) a.s. for all $\epsilon > 0$.

**Remark.** When $E|X_n|^2 = 1$, we have $\limsup_{n \to \infty} \frac{S_n}{(2n \log(\log(n)))^{1/2}} = 1$.

**Theorem 3.8.** Let $X_1, X_2, \cdots$ be i.i.d. mean 0, variance $\sigma^2$ random variables, and put $S_n := X_1 + \cdots + X_n$ for each $n \geq 1$. Then we have $\limsup_n S_n = \infty$ a.s. (U of Utah spring 2008 qualifying exam; may use CLT and Komolgorv's 0-1 law to prove it.)

**Theorem 3.9.** Let $X_1, X_2, \cdots$ be i.i.d. where $E[|X_n|] = \infty$. Then $\limsup_n \frac{|S_n|}{n} = \infty$ a.s. (Chung; hint: prove $P(|X_n| > n \epsilon \ i.o) = 1$, and then $P(|S_n| > n \epsilon \ i.o) = 1$.)

4 Random Series

**Theorem 4.1.** (Kolmogrov’s 3 series theorem) Let $X_1, X_2, \cdots$ be independent. Let $A > 0$ and $Y_n = X_n 1_{\{|X_n| \leq A\}}$. Then $\sum_{n=1}^{\infty} X_n$ converges a.s. if and only if all the following three conditions hold:

1. $\sum_{n=1}^{\infty} P(|X_n| > A) < \infty$.
2. $\sum_{n=1}^{\infty} E[Y_n] < \infty$.
Theorem 4.2. (Levy’s theorem) If $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of independent random variables, and $S_n := X_1 + X_2 + \cdots + X_n$ then TFAE: (Varadhan)

1. $S_n$ converges weakly to some r.v. $S$.
2. $S_n$ converges in pr. to some r.v. $S$.
3. $S_n$ converges a.s. to some r.v. $S$.

Remark. For (2) $\Rightarrow$ (3), we may refer to Theorem 5.3.4 of Chung.

5 Uniform Integrability

Theorem 5.1. The family $\{X_\alpha\}$ is uniformly integrable if and only if the following two conditions are satisfied: (Chung)

1. $\sup_\alpha E|X_\alpha| < \infty$.
2. For every $\epsilon > 0$, there exists some $\delta(\epsilon) > 0$ s.t. for any $P(E) < \delta(\epsilon)$, we have $\int_E |X_\alpha| dP < \epsilon$.

Theorem 5.2. Let $X_n \to X$ in dist, and for some $p > 0$ we have that $\sup_n E|X_n|^p < \infty$. We then have $E|X_n|^r \to E|X|^r$ for every $r < p$. (Chung: it provides us with a criterion to check uniform integrability; see Theorem 5.4.)

Theorem 5.3. If $\sup_n E|X_n|^p < \infty$ for some $p > 1$, then $\{X_n\}$ is u.i. (Consider $E[|X_n|; |X_n| > A] \leq (E|X_n|^p)^{1/p}(\frac{E|X_n|^p}{A^p})^{1/q}$)(Chung)

Theorem 5.4. If $\{X_n\}$ is dominated by some $Y \in L^p$, and converges in dist. to $X$, then $E|X_n|^p \to E|X|^p$. (Chung; to prove it, first show that $E|X|^p < \infty$, otherwise we may pick $A$ large so that $E[|X|^p \wedge A] > E|Y|^p$; also use the convergence $E[|X_n|^p \wedge A] \to E[|X|^p \wedge A]$ for every finite $A > 0$. This theorem can be also used to check uniform integrability; see Theorem 5.4.)

Theorem 5.5. Let $\sup_n |X_n| \in L^p$ and $X_n \to X$. Then $X \in L^p$ (by Fatou) and $X_n \to X$ in $L^p$ (by DCT). (Chung)

Theorem 5.6. Let $0 < p < \infty$, $X_n \in L^p$, and $X_n \to X$ in pr. Then
TFAE: (Chung)

1. $|X_n|^p$ is u.i.
2. $X_n \to X$ in $L^p$.
3. $E|X_n|^p \to E|X|^p < \infty$.

Theorem 5.7. Given $(\Omega, F, P)$, \{\[E[X|G] : G \text{ is a } \sigma\text{-algebra}, G \subset F\} is u.i. (Durrett; prove it with Theorem 5.1)

Theorem 5.8. (Dunford-Pettis) Let $(X, \Sigma, \mu)$ be any measure space and $A$ a subset of $L^1 = L^1(\mu)$. Then $A$ is uniformly integrable iff it is relatively compact in $L^1$ for the weak topology of $L^1$. (measure theory II by D.H.Fremlin)

6 Martingales

6.1 General submartingales

Theorem 6.1. (Producing submartingales with submartingales) If $\{X_n\}$ is a martingale and $\phi$ is convex, then $\phi(X)$ is a submartingale, provided that $\phi(X_n) \in L^1(P)$ for all $n$. If $X$ is a submartingale and $\phi$ is a nondecreasing convex function, then $\phi(X)$ is a submartingale, provided that $\phi(X_n) \in L^1(P)$ for all $n$. (Davar)

Theorem 6.2 (Doob’s Decomposition) Any submartingale $\{X_n\}$ can be written as $X_n = Y_n + Z_n$, where $\{Y_n\}$ is a martingale, and $\{Z_n\}$ is a non-negative predictable a.s. increasing process with $Z_n \in L^1(P)$ for all $n$. (Davar)

Remark. By the inequalities $EZ_n \leq E|X_n| - EY_1$ and $E|Y_n| \leq E|X_n| + EZ_n$, we have that $\{X_n\}$ is $L^1$-bounded if and only if both $\{Y_n\}$ and $\{Z_n\}$ are $L^1$-bounded, and since $E[\lim_n Z_n] < \infty$, we have that $\{X_n\}$ is u.i. if and only if both $\{Y_n\}$ and $\{Z_n\}$ are u.i. (Chung)

Theorem 6.3. (Doob’s maximal inequality) If $\{X_n\}$ is a submartingale, then for all $\lambda > 0$ and $n \geq 1$, and we have (Davar)

1. $\lambda P(\max_{1 \leq j \leq n} X_j \geq \lambda) \leq E[X_n; \max_{1 \leq j \leq n} X_j \geq \lambda] \leq E[X_n^+]$.
2. $\lambda P(\min_{1 \leq j \leq n} X_j \leq -\lambda) \leq E[X_n^+] - EX_1$.
3. Combine (1) and (2), we have $\lambda P(\max_{1 \leq j \leq n} |X_j| \geq \lambda) \leq 2E|X_n| - EX_1$.
Remarks. (1) If \( \{X_n\} \) is a martingale, then
\[
P(\max_{1 \leq j \leq n} |X_j| \geq \lambda) \leq \frac{E|X_n|^p}{\lambda^p}
\]
for \( p \geq 1 \) and \( \lambda > 0 \). This is because \( |X_n|^p \) is a submartingale. (2) We may take \( p = 2 \) and replace \( X_n \) with \( S_n - ES_n \) in (1) to obtain Kolmogrov’s maximal inequality.

### 6.2 Good submartingales (a.s. convergence)

**Theorem 6.4 (The Martingale Convergence Theorem)** Let \( \{X_n\} \) be an \( L^1 \)-bounded submartingale. Then \( \{X_n\} \) converges a.s. to a finite limit.

*(Chung)*

Remarks. (1) As a corollary, every nonnegative supermartingale and nonpositive submartingale converges a.s. to a finite limit. (2) It suffices to assume \( \sup_n E|X_n| < \infty \) (\( \iff \sup_n E|X_n| < \infty \)). This is due to \( E|X_n| = 2EX_n^+ - EX_n \leq 2EX_n^+ - EX_1 \). (3) It does not hold conversely. For example, we may keep throwing a fair coin and bet \( 10^n \) dollars on the \( n \)-th round. We stop when our money becomes negative.

**Theorem 6.5 (Krickeberg’s Decomposition)** Suppose \( \{X_n\} \) is a submartingale which is bounded in \( L^1(P) \). Then we can write \( X_n = Y_n - Z_n \), where \( \{Y_n\} \) is a martingale and \( \{Z_n\} \) is a non-negative supermartingale.

Remarks. (1) By the inequalities
\[
E|Z_n| = |EZ_n| \leq E|X_n| + |EY_1| \quad \text{and} \quad E|Y_n| \leq E|X_n| + E|Z_n| = E|X_n| + EZ_n,
\]
we have that both \( \{Y_n\} \) and \( \{Z_n\} \) are \( L^1 \)-bounded. (Davar) (2) Unlike Theorem 6.2, when \( \{X_n\} \) is u.i., it does not imply that both \( \{Y_n\} \) and \( \{Z_n\} \) are u.i.. To see this, we may simply assume that \( \{X_n\} \) is a nonpositive submartingale which is not u.i..

### 6.3 Better submartingales (\( L^1 \) convergence or \( L^p \) convergence, \( p > 1 \))

**Theorem 6.6** For a submartingale, TFAE: (Durrett)

1. It is u.i.
2. It converges a.s. and in \( L^1 \).
3. It converges in \( L^1 \).
Remark. Say, it converges to \( X_\infty \), which can be regarded as “the last term” of the original submartingale.

**Theorem 6.7** For a martingale, TFAE: (Durrett)

\begin{enumerate}
  
  \item It is u.i.
  \item It converges a.s. and in \( L^1 \).
  \item It converges in \( L^1 \).
  \item There exists \( X \) with \( E|X| < \infty \) so that \( X_n = E[X|F_n] \).
\end{enumerate}

**Theorem 6.8** \( \text{Let } F_n \uparrow F, \text{ then } E[X|F_n] \to E[X|F] \text{ a.s. and in } L^1. \) (Durrett)

Remark. For a generalization, see Chung Theorem 9.4.8.

**Theorem 6.9 (L^p maximum inequality)** If \( X_n \) is a nonnegative submartingale s.t. \( \sup_n E[X_n^p] < \infty \), then for \( 1 < p < \infty \),
\[
E[(\max_{1 \leq j \leq n} X_j)^p] \leq \left( \frac{p}{1-p} \right)^p E[X_n]^p.
\] (Davar; Chung)

Remark. That is, if a nonnegative submartingale if \( L^p \)-bounded, then it is bounded by a \( L^p \) function. This inequality also holds if we replace \( X_n \) with \( |X_n| \) and submartingales with martingales.

**Theorem 6.10 (L^1 maximum inequality)** If \( X_n \) is a submartingale, then
\[
E[\max_{1 \leq j \leq n} X_j^+] \leq \frac{1}{1-1/e}(1 + E[X_n^+ \max(\log(X_n^+), 0)]).
\] (Durrett)

Remark. By Theorem 5.5 and this theorem, we have another sufficient condition for \( L^1 \) convergence martingales.

**Theorem 6.11 (L^p convergence theorem for martingales or non-negative submartingales)** If \( X_n \) is a nonnegative submartingale or a martingale with \( \sup_n E|X_n|^p < \infty \), \( p > 1 \), then \( X_n \to X \) a.s and in \( L^p \). (Durrett)

Remark. For a proof, see Durrett Theorem 4.4.5 or Varadhan Theorem 5.6; we may also use (a) Theorem 5.5 + 6.9 (b) Theorem 5.4 + 5.6 + 6.9 to prove it.

**Theorem 6.12 (L^p convergence theorem for submartingales)** If \( X_n \) is a submartingale with \( \sup_n E|X_n|^p < \infty \), \( p > 1 \), then \( X_n \to X \) a.s and in \( L^r \) for every \( 1 \leq r < p \). (A combination of various theorems.)
Remarks. We may use Theorem 5.2 + 5.6 + 6.4 (martingale convergence theorem) to prove it. (Here we don’t use theorem 6.9)

7 Stopping times

Theorem 7.1 (“generalized Optional Stopping Theorem”) If $L \leq M$ are stopping times and $\{Y_{M \wedge n}\}$ is a uniformly integrable submartingale, then $EY_L \leq EY_M$ and $Y_L \leq E[Y_M | F_L]$.(Durrett)

Remarks. (1) If $L \leq M$ are two bounded stopping times (Say, $L \leq M \leq k$), then $|Y_{M \wedge n}| \leq |Y_1| + \cdots + |Y_k|$, which means $\{Y_{M \wedge n}\}$ is u.i.. Such case is usually called Optional Stopping Theorem. (2) There’re two sufficient conditions for $\{Y_{M \wedge n}\}$ to be u.i.(Durrett): (a) $\{Y_{M \wedge n}\}$ is a u.i. submartingale. (b) $E|Y_M| < \infty$ and $\{Y_n 1_{\{M > n\}}\}$ is u.i.

8 Truncation

Theorem 8.1 ($L^1$-Truncation) Let $X_1, X_2, \cdots$ be i.i.d with $E|X_1| < \infty$, and let $Y_n = X_n 1_{|X_n| \leq n}$. Then $P(X_n \neq Y_n \ i.o.) = 0$.

Theorem 8.2 ($L^p$-Truncation, $p > 1$) Let $X_1, X_2, \cdots$ be i.i.d with $E[|X_1|^p] < \infty$, and let $Y_n = X_n 1_{|X_n| \leq n^{1/p}}$. Then $P(X_n \neq Y_n \ i.o.) = 0$.

9 Examples of Martingales

Example 9.1 Let $X_1, X_2, \cdots$ be independent and $EX_n = 0$ for all $n \geq 1$. Then $S_n := X_1 + \cdots + X_n$ is a martingale.

Example 9.2 Let $X_1, X_2, \cdots$ be independent, $EX_n = 1$ for all $n \geq 1$, and $Y_n := X_1 \times \cdots \times X_n$. Then $\{Y_n\}$ is a martingale.

Example 9.3 (Related to densities- Likelihood ratios) Suppose $f$ and $g$ are two strictly positive probability density functions on $\mathbb{R}$. If $\{X_i\}_{i=1}^\infty$ are i.i.d. random variables with probability density $f$, then $\Pi_{j=1}^n [g(X_j)/f(X_j)]$ defines a mean-one martingale. (Davar)
Example 9.4 Let $X_1, X_2, \cdots$ be independent, $E|X_n|^2 < \infty$ and $EX_n = 0$ for all $n \geq 1$, and $S_n := X_1 + \cdots + X_n$. Then $\{S_n^2 - \sum_{j=1}^{\infty} E[X_j^2]\}$ is a martingale.

Example 9.5 If $X_n$ and $Y_n$ are submartingales w.r.t. $F_n$ then $X_n \lor Y_n$ also is. (Durrett)

Example 9.6 Suppose $E|X_n| < \infty$ for all $n \geq 1$, and $E[X_{n+1}|X_1, \cdots, X_n] = \frac{X_1 + \cdots + X_n}{n}$. Then $\{\frac{X_1 + \cdots + X_n}{n}, \sigma(X_1, \cdots, X_n)\}$ is a martingale. (Chung)

Example 9.7 (Martingales induced from Markov chains.) Let $\{X_n\}$ be a Markov chain and $f = Pf$ (That is, $f$ is $p$-harmonic). Then $f(X_n)$ is a martingale. Note that if $\{X_n\}$ is recurrent then $f$ is trivial. (Karlin)

Example 9.8 (Martingales induced from Brownian Motions(I)) $\exp(\theta B_t - (\theta^2 t/2))$ is a martingale for all $\theta \in \mathbb{R}$. (Durrett)

Example 9.9 (Martingales induced from Brownian Motions(II)) $B_t, B^2_t - t, B^3_t - 3tB_t, B^4_t - 6tB^2_t + 3t^2, \cdots$ are all martingales. (Durrett)

Remark. When we’re given a 1-parameter family of martingales, we may differentiate with respect to that parameter and see what happens.

Example 9.10 Let $Y_1, Y_2, \cdots$ be independent random variables. Define $X_0 = 0, X_n = \sum_{j=1}^{n} Y_j - E[Y_j|Y_1, \cdots, Y_{j-1}]$ for $n \geq 1$. Then $\{X_n\}$ is a martingale. (A probability path; Resnick)

10 References

Main reference: Chung, Durrett, Davar.

Other references: Varadhan, Shiryaev, Resnick, Karatzas, Karlin, Breiman.