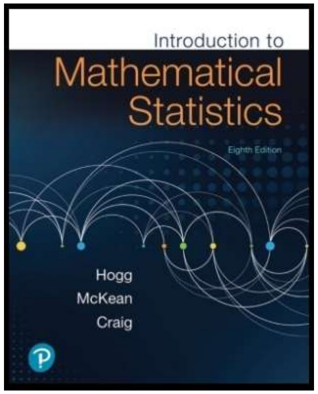
Mathematical Statistics 1

Chapter 5. Consistency and Limiting Distributions

5.2. Convergence in Distribution—Proofs of Theorems



Theorem 5.2.1

Theorem 5.2.1. If sequence of random variables (X_n) converges to X in probability, then (X_n) converges to X in distribution.

Proof. Let x be a point of continuity of the cumulative distribution function $F_X(x)$ and let $\varepsilon > 0$. Then

$$\begin{aligned}
F_{X_n}(x) &= P(X_n \le x) \\
&= P((X_n \le x) \cap (\{|X_n - X| < \varepsilon\} \cup \{|X_n - X| \ge \varepsilon\}) \\
&= P(X_n \le x) \cap \{|X_n - X| < \varepsilon\} + P(X_n \le x) \cap \{|X_n - X| \ge \varepsilon\}) \\
&\le P(X_n \le x) \cap \{|X_n - X| < \varepsilon\} + P(|X_n - X| \ge \varepsilon) \\
&= P(X_n \le x) \cap \{-\varepsilon < X_n - X < \varepsilon\}) \\
&\quad + P((X_n \le x) \cap \{|X_n - X| \ge \varepsilon\}) \\
&\le P((X_n \le x) \cap \{X - X_n < \varepsilon\}) + P(|X_n - X| \ge \varepsilon) \\
&\le P(X \le x + \varepsilon) + P(|X_n - X| \ge \varepsilon) \text{ since } X_n \le x \text{ and} \\
&\quad X - X_n < \varepsilon \text{ together imply } X \le X_n + \varepsilon \le x + \varepsilon.
\end{aligned}$$

Theorem 5.2.1 (continued 1)

Proof (continued). Since $X_n \xrightarrow{P} X$ then be definition $\lim_{n\to\infty} P(|X_n - | \ge \varepsilon) = 0$, so

$$\begin{aligned}
\overline{\lim_{n\to\infty}} F_n &\le \overline{\lim_{n\to\infty}} P\left(|X \le x + \varepsilon) + P(|X_n - X| \ge \varepsilon)\right) \\
&= \overline{\lim_{n\to\infty}} P(X \le x + \varepsilon) + \overline{\lim_{n\to\infty}} P(|X_n - X| \ge \varepsilon) \\
&= P(X \le x + \varepsilon) + 0 = F_X(x + \varepsilon). \qquad (5.2.5)
\end{aligned}$$

Similarly,

$$\begin{aligned}
P(X_n > x) &= P((X_n > x) \cap (\{|X_n - X| < \varepsilon\} \cup \{|X_n - X| \ge \varepsilon\}) \\
&= P((X_n > x) \cap \{|X_n - X| < \varepsilon\}) \\
&\quad + P((X_n > x) \cap \{|X_n - X| \ge \varepsilon\}) \\
&\le P((X_n > x) \cap \{|X_n - X| < \varepsilon\}) + P(|X_n - X| \ge \varepsilon) \\
&= P((X_n > x) \cap \{-\varepsilon < X_n - X < \varepsilon\}) \\
&\quad + P((X_n > x) \cap \{|X_n - X| \ge \varepsilon\}) \dots
\end{aligned}$$

Theorem 5.2.1 (continued 2)

Proof (continued). ...

$$\begin{aligned}
P(X_n > x) &\le P((X_n > x) \cap \{-\varepsilon < X_n - X < \varepsilon\}) \\
&\quad + P((X_n > x) \cap \{|X_n - X| \ge \varepsilon\}) \\
&\le P((X_n > x) \cap \{-\varepsilon < X - X_n\}) + P(|X_n - X| \ge \varepsilon) \\
&\le P(X \le x - \varepsilon) + P(|X_n - X| \ge \varepsilon) \text{ since } X_n \le x \text{ and} \\
&\quad -\varepsilon < X - X_n \text{ together imply } X > X_n - \varepsilon > x - \varepsilon.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\underline{\lim_{n\to\infty}} (1 - F_{X_n}(x)) &= 1 + \underline{\lim_{n\to\infty}} (-F_{X_n}(x)) = 1 - \overline{\lim_{n\to\infty}} F_{X_n}(x) \\
&= \overline{\lim_{n\to\infty}} (1 - F_{X_n}(x)) = \overline{\lim_{n\to\infty}} P(X_n > x) \\
&\le \overline{\lim_{n\to\infty}} (P(X \ge x - \varepsilon) + P(|X_n - X| \ge \varepsilon)) \\
&= \overline{\lim_{n\to\infty}} P(X \ge x - \varepsilon) + \overline{\lim_{n\to\infty}} P(|X_n - X| \ge \varepsilon) \\
&= P(X \ge x - \varepsilon) + 0 = 1 - F_X(x - \varepsilon). \qquad (5.2.6)
\end{aligned}$$

Theorem 5.2.1 (continued 3)

Theorem 5.2.1. If sequence of random variables (X_n) converges to X in probability, then (X_n) converges to X in distribution.

Proof (continued). Combining (5.2.5) and (5.2.6), we get

$$F_X(x-\varepsilon) \leq \underline{\lim}_{n\to\infty} F_{X_n}(x) \leq \overline{\lim}_{n\to\infty} F_{X_n}(x) \leq F_X(x+\varepsilon).$$

Since $\varepsilon > 0$ is arbitrary and since x is a point of continuity of F_X then we must have

$$\lim_{n\to\infty} F_{X_n}(x) \text{ exists and } \lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

Since x is an arbitrary element of $C(F_X)$, then X_n converges in distribution to X (by definition), as claimed. □

Theorem 5.2.2

Theorem 5.2.2. If sequence of random variables (X_n) converges to constant b in distribution, then (X_n) converges to b in probability.

Proof. Let $\varepsilon > 0$. Since $X_n \xrightarrow{D} b$ by hypothesis then $\lim_{n\to\infty} F_{X_n}(x) = F_X(x) = b$ for all $x \in C(F_X)$. So

$$\lim_{n\to\infty} P(|X_n - b| \leq \varepsilon) = \lim_{n\to\infty} P(-\varepsilon \leq X_n - b \leq \varepsilon)$$

$$\begin{aligned}
&= \lim_{n\to\infty} P(b-\varepsilon \leq X_n \leq b+\varepsilon) = \lim_{n\to\infty}(P(X_n \leq b+\varepsilon) - P(X_n < b-\varepsilon)) \\
&= \lim_{n\to\infty}(P(X_n \leq b+\varepsilon) - P(X_n \leq b-\varepsilon) + P(X_n - b - \varepsilon)) \\
&= \lim_{n\to\infty} F_{X_n}(b+\varepsilon) = \lim_{n\to\infty} F_{X_n}(b-\varepsilon) + \lim_{n\to\infty} P(X_n = b-\varepsilon) \\
&= 1 - 0 + \lim_{n\to\infty} P(X_n = b-\varepsilon) \text{ since } \lim_{n\to\infty} F_{X_n}(x) = b \\
&= 1 \text{ since a limit of probabilities must be at most 1.}
\end{aligned}$$

(Notice that we must have $\lim_{n\to\infty} P(X = b-\varepsilon) = 0$). Therefore, by definition, $X_n \xrightarrow{P} b$, as claimed. □

Theorem 5.2.6

Theorem 5.2.6. Let (X_n) be a sequence of random variables and let X be a random variable. If $X_n \to X$ in distribution, then (X_n) is bounded in probability.

Proof. Consider the cumulative distribution functions F_{X_n} and F_X. Let $\varepsilon > 0$ be given and choose η so that η and $-\eta$ are continuity points of F_X and $P(|X| \leq \eta) \geq 1-\varepsilon$, as described in equation (5.2.7). Now

$$\begin{aligned}
\lim_{n\to\infty} P(|X_n| \leq \eta) &= \lim_{n\to\infty}(F_{X_n}(\eta) - F_{X_n}(-\eta) + P(X_n = -\eta)) \text{ by (5.2.7)} \\
&\geq \lim_{n\to\infty} F_{X_n}(\eta) - \lim_{n\to\infty} F_{X_n}(-\eta) + 0 \\
&= F_X(\eta) - F_X(-\eta) \text{ since } X_n \xrightarrow{D} X \\
&> 1-\varepsilon \text{ by (5.2.7).}
\end{aligned}$$

So with $B_\varepsilon = \eta$ there exists $N_\varepsilon \in \mathbb{N}$ such that for all $\eta \geq N_\varepsilon$ we have $P(|X_n| \leq X_\varepsilon) \geq 1-\varepsilon$ and so (X_n) is bounded in probability, as claimed. □

Theorem 5.2.7

Theorem 5.2.7. Let (X_n) be a sequence of random variables which is bounded in probability and let (Y_n) be a sequence of random variables that converges to 0 in probability. Then $X_n Y_n \xrightarrow{P} 0$.

Proof. Since $\varepsilon > 0$. Since (X_n) is bounded in probability by hypothesis then there exists $B_\varepsilon > 0$ and $N_\varepsilon \in \mathbb{N}$ such that

$$n \geq N_\varepsilon \text{ implies } P(|X_n| \leq B_\varepsilon) \geq 1-\varepsilon.$$

Then

$$\overline{\lim}_{n\to\infty} P(|X_nY_n| \geq \varepsilon) = \overline{\lim}_{n\to\infty} P\left((|X_nY_N| \geq \varepsilon) \cap (|X_n| \leq B_\varepsilon) \cup (|X_n| > B_\varepsilon)\right)$$

$$= \overline{\lim}_{n\to\infty}(P((|X_nY_n| \geq \varepsilon) \cap (|X_n| \leq B_\varepsilon)) + P((|X_nY_n| \geq \varepsilon) \cap (|X_n \ > B_\varepsilon))$$

$$\leq \overline{\lim}_{n\to\infty} P((|X_nY_n| \geq \varepsilon) \cap (|X_n| \leq B_\varepsilon)) + \overline{\lim}_{n\to\infty} P((|X_nY_n| \geq \varepsilon) \cap (|X_n| > B_\varepsilon))$$

Theorem 5.2.7 (continued)

Proof (continued). ...

$$\le \overline{\lim_{n\to\infty}} P((|X_nY_n| \ge \varepsilon)\cap(|X_n| \le B_\varepsilon)) + \overline{\lim_{n\to\infty}} P((|X_nY_n| \ge \varepsilon)\cap(|X_n| > B_\varepsilon))$$

$$\begin{aligned} &\le \overline{\lim_{n\to\infty}} P(|Y_n| \ge \varepsilon/B_\varepsilon) + \overline{\lim_{n\to\infty}} P(|X_n| > B_\varepsilon) \text{ since } |X_n| > B_\varepsilon \text{ implies} \\ &\quad 1/|X_n| < 1/B_\varepsilon \text{ so that } P(|X_nY_n| \ge \varepsilon) = P(|Y_n| \ge \varepsilon/|X_n|) \\ &\quad \ge P(|Y_n| \ge \varepsilon/B_\varepsilon) \text{ since } \varepsilon/B_\varepsilon > \varepsilon/|X_n| \\ &\le \overline{\lim_{n\to\infty}} P(|Y_n| \ge \varepsilon/B_\varepsilon \text{ since } P(|X_n| \le B_\varepsilon) \ge 1-\varepsilon \\ &\quad \text{implies } \varepsilon \ge 1 - P(|X_n| \le B_\varepsilon) = P(|X_n| > B_\varepsilon) \\ &= 0 + \varepsilon = \varepsilon \text{ since } Y_N \xrightarrow{P} 0. \end{aligned}$$

So $\overline{\lim}_{n\to\infty} P(|X_nY_n| \ge \varepsilon) = 0$ and, since $P(|X_nY_n| \ge \varepsilon) \ge 0$, then $\underline{\lim}_{n\to\infty} P(|X_nY_n| \ge \varepsilon) = 0$. Hence $\lim_{n\to\infty} P(|X_nY_n| \ge \varepsilon) = 0$ so that, by definition of convergence in probability, $X_nY_n \xrightarrow{P} 0$, as claimed. □

Theorem 5.2.8

Theorem 5.2.8. Suppose sequence of random variables (Y_n) is bounded in probability. Suppose $X_n = o_p(Y_n)$. Then $X_n \xrightarrow{P} 0$ as $n \to \infty$.

Proof. Let $\varepsilon > 0$ and let $\varepsilon' > 0$. Since (Y_n) is bounded in probability by hypothesis, then (by definition) there events $N_{\varepsilon'} \in \mathbb{N}$ and $B_{\varepsilon'} > 0$ such that if $n \ge N_{\varepsilon'}$ then $P(|Y_n| \le B_{\varepsilon'}) \ge a - \varepsilon'/2$. Also, since $X_n = o_p(Y_n)$ by hypothesis then (by definition) $X_n/Y_n \xrightarrow{P} 0$ as $n \to \infty$; that is, $\lim_{n\to\infty} P(|S_n/Y_n| \ge \varepsilon) = 0$. So there is $N_1 \in \mathbb{N}$ such that for all $n \ge N_1$ we have $P(|X_n/Y_n| \ge \varepsilon/B_{\varepsilon'}) < \varepsilon'/2$. Let $N = \max\{N_{\varepsilon'}, N_1\}$. Then for $n \ge N$ we have

$$\begin{aligned} P(|X_n| \ge \varepsilon) &= P((|X_n| \ge \varepsilon) \cap ((|Y_n| \le B_{\varepsilon'}) \cap (|Y_n| > B_{\varepsilon'})) \\ &= P((|X_n| \ge \varepsilon) \cap (|Y_n| \le B_{\varepsilon'})) + P((|X_n| \ge \varepsilon) \cap (|Y_n| > B_{\varepsilon'})) \\ &\le P(|X_n|/Y_n| \ge \varepsilon/B_{\varepsilon'}) + P(|Y_n| > B_{\varepsilon'}) \text{ since } |X_n| \ge \varepsilon \\ &\quad \text{and } |Y_n| \le B_{\varepsilon'} \text{ implies } |X_n/Y_n| \ge \varepsilon/B_{\varepsilon'} \end{aligned}$$

Theorem 5.2.8 (continued)

Theorem 5.2.8. Suppose sequence of random variables (Y_n) is bounded in probability. Suppose $X_n = o_p(Y_n)$. Then $X_n \xrightarrow{P} 0$ as $n \to \infty$.

Proof (continued). ...

$$\begin{aligned} P(|X_n| \ge \varepsilon) &\le P(|X_n|/Y_n| \ge \varepsilon/B_{\varepsilon'}) + P(|Y_n| > B_{\varepsilon'}) \\ &< \varepsilon'/2 + (1 - P(|Y_n| \le B_{\varepsilon'})) < \varepsilon'/2 + (1 - (1 - \varepsilon'/2)) \\ &= \varepsilon'/2 + \varepsilon'/2 = \varepsilon'. \end{aligned}$$

Since $\varepsilon' > 0$ is arbitrary, then $\lim_{n\to\infty} P(|X_n| \ge \varepsilon) = 0$ and so $X_n \xrightarrow{P} 0$ as $n \to \infty$, as claimed. □

Theorem 5.2.9

Theorem 5.2.9. Let (X_n) be a sequence of random variables which that $\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2)$. Suppose the function $g(x)$ is differentiable at θ and $g'(\theta) \ne 0$. Then

$$\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2(g'(\theta))^2).$$

Proof. By "Theorem 5.2.A. A General Mean Value Theorem" we have

$$g(X_n) = g(\theta) + g'(\theta)(X_n - \theta) + o_p(|X_n - \theta|),$$

or

$$g(X_n) - g(\theta) = g'(\theta)(X_n - \theta) + o_p(|X_n - \theta|)$$

or

$$\sqrt{n}g(X_n) - \sqrt{n}g(\theta) = g'(\theta)\sqrt{n}(X_n - \theta) + \sqrt{n}o_p(|X_n - \theta|).$$

Now $Y_n = o_p(X_n)$ means $Y_n/X_n \xrightarrow{P} 0$ as $n \to \infty$ (or $\lim_{n\to\infty} P(|Y_n/X_n| \ge \varepsilon) = 0$), so $Y_n = o_p(\sqrt{n}X_n)$ since $\lim_{n\to\infty} P(|Y_n/(\sqrt{n}X_n)| \ge \varepsilon) = 0$.

Theorem 5.2.9 (continued)

Proof (continued). Hence

$$\sqrt{n}g(X_n) - \sqrt{n}g(\theta) = g'(\theta)\sqrt{n}(X_n - \theta) + o_p(\sqrt{n}|X_n - \theta|).$$

Since $\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2)$, then by Theorem 5.2.6 $\sqrt{n}(X_n - \theta)$ is bounded in probability and hence so is $\sqrt{n}|X_n - \theta|$. By Theorem 5.2.8 (since $\sqrt{n}|X_n = \theta|$ is bounded in probability) then $o_p(\sqrt{n}|X_n - \theta|) \xrightarrow{P} 0$ as $n \to \infty$ which, by Theorem 5.2.1. Therefore,

$$\sqrt{n}g(X_n) - \sqrt{n}g(\theta) = \sqrt{n}(g(X_n) - g(\theta)) = g'(\theta)\sqrt{n}(X_n - \theta) + o_p(\sqrt{n}|X_n - \theta|).$$

That is, $\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2)$ and $o_p(\sqrt{n}|X_n - \theta|) \xrightarrow{P} 0$, so by Theorem 2.5.2 (Slutsky's Theorem),

$$\sqrt{n}g(X_n) - \sqrt{n}g(\theta) = \sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{D} g'(\theta)N(0, \sigma^2) + 0.$$

Now $g'(\theta)N(0, \sigma^2) = N(0, \sigma^2(g'(\theta))^2)$ by Theorem 2.4.2, therefore $\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2(g'(\theta))^2)$, as claimed. □