Mathematical Statistics 1

Chapter 5. Consistency and Limiting Distributions

5.4. Extensions to Multivariate Distributions—Proofs of Theorems

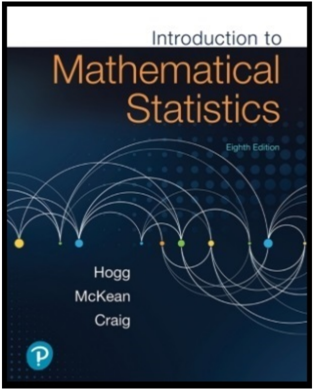

Table of contents

1. Lemma 5.3.1
2. Theorem 5.4.1
3. Theorem 5.4.4. Multivariate Central Limit Theorem

Lemma 5.4.1

Lemma 5.4.1. Let $\mathbf{v}' = (v_1, v_2, \ldots, v_p) \in \mathbb{R}^p$. Then $|v_j| \leq \|\mathbf{v}\| \leq \sum_{i=1}^p |v_i|$ for all $j = 1, 2, \ldots, p$.

Proof. For all $j = 1, 2, \ldots, p$ we have $v_j \leq \sum_{i=1}^p v_i^2 = \|\mathbf{v}\|^2$ and so (taking square roots) $|v_j| \leq \|\mathbf{v}\|$. Next, by the Triangle Inequality,

$$\|\mathbf{v}\| = \left\{ \sum_{i=1}^p v_i \mathbf{e}_i \right\| \leq \sum_{i=1}^p \|v_i \mathbf{e}_i\| = \sum_{i=1}^p |v_i|$$

(since $\|\mathbf{e}_i\| = 1$ for each i). Therefore, $|v_j| \leq \|\mathbf{v}\| \leq \sum_{i=1}^p |v_i|$, as claimed. □

Lemma 5.4.1

Lemma 5.4.1. Let $\mathbf{v}' = (v_1, v_2, \ldots, v_p) \in \mathbb{R}^p$. Then $|v_j| \le \|\mathbf{v}\| \le \sum_{i=1}^p |v_i|$ for all $j = 1, 2, \ldots, p$.

Proof. For all $j = 1, 2, \ldots, p$ we have $v_j \le \sum_{i=1}^p v_i^2 = \|\mathbf{v}\|^2$ and so (taking square roots) $|v_j| \le \|\mathbf{v}\|$. Next, by the Triangle Inequality,

$$\|\mathbf{v}\| = \left\{ \sum_{i=1}^p v_i \mathbf{e}_i \right\| \le \sum_{i=1}^p \|v_i \mathbf{e}_i\| = \sum_{i=1}^p |v_i|$$

(since $\|\mathbf{e}_i\| = 1$ for each i). Therefore, $|v_j| \le \|\mathbf{v}\| \le \sum_{i=1}^p |v_i|$, as claimed. □

Theorem 5.4.1

Theorem 5.4.1. Let $\{\mathbf{X}_n\}$ be a sequence of p-dimensional vectors and let $\mathbf{X}$ be a random vector, all defined on the same sample space. Then

$$\mathbf{X}_n \xrightarrow{P} \mathbf{X} \text{ if and only if } X_{nj} \xrightarrow{P} X_j \text{ for all } j = 1, 2, \ldots, p.$$

Proof. Let $\varepsilon > 0$.

Suppose $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$. For any $j = 1, 2, \ldots, p$, if we have $\varepsilon \leq |X_{nj} - X_j|$ then

$$\varepsilon \leq |X_{nj} = X_j| \leq \|\mathbf{X}_n - \mathbf{X}\| \text{ by Lemma 5.4.1.}$$

Hence $\overline{\lim}_{n\to\infty} P(|X_{nj} - X_j| \geq \varepsilon) \leq \overline{\lim}_{n\to\infty} P(\|\mathbf{X}_n = \mathbf{X}\| \geq \varepsilon) = 0$ since $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$. Since $P(|X_{nj} - X_j| \geq \varepsilon) \geq 0$ for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} P(|X_{nj} - X_j| \geq \varepsilon) = 0$ and hence $X_{nj} \xrightarrow{P} X_j$ for all $j = 1, 2, \ldots, p$, as claimed.

Theorem 5.4.1

Theorem 5.4.1. Let $\{\mathbf{X}_n\}$ be a sequence of p-dimensional vectors and let $\mathbf{X}$ be a random vector, all defined on the same sample space. Then

$$\mathbf{X}_n \xrightarrow{P} \mathbf{X} \text{ if and only if } X_{nj} \xrightarrow{P} X_j \text{ for all } j = 1, 2, \ldots, p.$$

Proof. Let $\varepsilon > 0$.

Suppose $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$. For any $j = 1, 2, \ldots, p$, if we have $\varepsilon \leq |X_{nj} - X_j|$ then

$$\varepsilon \leq |X_{nj} = X_j| \leq \|\mathbf{X}_n - \mathbf{X}\| \text{ by Lemma 5.4.1.}$$

Hence $\overline{\lim}_{n\to\infty} P(|X_{nj} - X_j| \geq \varepsilon) \leq \overline{\lim}_{n\to\infty} P(\|\mathbf{X}_n = \mathbf{X}\| \geq \varepsilon) = 0$ since $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$. Since $P(|X_{nj} - X_j| \geq \varepsilon) \geq 0$ for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} P(|X_{nj} - X_j| \geq \varepsilon) = 0$ and hence $X_{nj} \xrightarrow{P} X_j$ for all $j = 1, 2, \ldots, p$, as claimed.

Theorem 5.4.1 (continued)

Proof (continued). Conversely, suppose $X_{nj} \xrightarrow{P} X_j$ for all $j = 1, 2, \ldots, p$. If we have $\varepsilon \le \|\mathbf{X}_n - \mathbf{X}\|$ then

$$\varepsilon \le \|\mathbf{X}_n - \mathbf{X}\| \le \sum_{i=1}^{p} |X_{nj} - X_j| \text{ by Lemma 5.4.1.}$$

Hence

$$\begin{aligned}
\overline{\lim_{n\to\infty}} P(\|\mathbf{X}_n - \mathbf{X}\| \ge \varepsilon) &\le \overline{\lim_{n\to\infty}} P\left(\sum_{j=1}^{p} |X_{nj} - X_j| \ge \varepsilon\right) \\
&\le \sum_{j=1}^{p} \overline{\lim_{n\to\infty}} P(|X_{nj} - X_j| \ge \varepsilon/p) \\
&= 0 \text{ since } X_{nj} \xrightarrow{P} X_j \text{ for } j = 1, 2, \ldots, p \\
&\quad \text{ by hypothesis.}
\end{aligned}$$

Since $P(\|\mathbf{X}_n - \mathbf{X}\| \ge \varepsilon) \ge 0$ then $\lim_{n\to\infty} P(\|\mathbf{X}_n - \mathbf{X}\| \ge \varepsilon) = 0$ and hence $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$, as claimed. □

Theorem 5.4.4

Theorem 5.4.4. Multivariate Central Limit Theorem.
Let $(\mathbf{X}_n)$ be a sequence of independent and identically distributed ("iid") random vectors with common mean vector μ and variance-covariance matrix $\mathbf{\Sigma}$ which is positive definite. Assume that the common moment generating function $M(\mathbf{t})$ exists in an open neighborhood of $\mathbf{0}$. Let

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\mathbf{X}_i - \mu) = \sqrt{n}(\sqrt{\mathbf{X}} - \mu).$$

Then $\mathbf{Y}_n$ converges in distribution to a $N_p(\mathbf{0}, \mathbf{\Sigma})$ distribution.

Proof. Let $\mathbf{t} \in \mathbb{R}^p$ be in the neighborhood of $\mathbf{0}$ on which $M(\mathbf{t})$ exists. The moment generating function of $\mathbf{Y}_n$ is

$$M_n(\mathbf{t}) = E\left[\exp\left(\mathbf{t}' \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\mathbf{X}_i - \mu)\right)\right] \cdots$$

Theorem 5.4.4

Theorem 5.4.4. Multivariate Central Limit Theorem.
Let $(\mathbf{X}_n)$ be a sequence of independent and identically distributed ("iid") random vectors with common mean vector μ and variance-covariance matrix $\mathbf{\Sigma}$ which is positive definite. Assume that the common moment generating function $M(\mathbf{t})$ exists in an open neighborhood of $\mathbf{0}$. Let

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\mathbf{X}_i - \mu) = \sqrt{n}(\sqrt{\mathbf{X}} - \mu).$$

Then $\mathbf{Y}_n$ converges in distribution to a $N_p(\mathbf{0}, \mathbf{\Sigma})$ distribution.

Proof. Let $\mathbf{t} \in \mathbb{R}^p$ be in the neighborhood of $\mathbf{0}$ on which $M(\mathbf{t})$ exists. The moment generating function of $\mathbf{Y}_n$ is

$$M_n(\mathbf{t}) = E\left[\exp\left(\mathbf{t}' \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\mathbf{X}_i - \boldsymbol{\mu})\right)\right] \ldots$$

Theorem 5.4.4 (continued 1)

Proof (continued). ...

$$M_n(\mathbf{t}) = E\left[\exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{t}'(\mathbf{X}_i - \boldsymbol{\mu})\right)\right] = E\left[\exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{W}_i\right)\right]$$

where $\mathbf{W}_i = \mathbf{t}'(\mathbf{X}_i = \boldsymbol{\mu})$. Note that $\mathbf{W} - 1, \mathbf{W}_2, \ldots, \mathbf{W}_n$ are iid (since $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are) with mean $\mathbf{0}$ and variance

$$\mathrm{Var}(W_i) = (W_i - 0)^2 = \mathbf{t}'(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{t}'(\mathbf{X}_i - \boldsymbol{\mu}))'$$

$$= \mathbf{t}'(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i' - \boldsymbol{\mu}')\mathbf{t} = \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}.$$

So by the Central Limit Theorem (Theorem 5.3.1) and Note 5.3.A,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} W_i \xrightarrow{D} N(0, \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}). \qquad (*)$$

Theorem 5.4.4 (continued 2)

Proof (continued). Since the moment generating function of $\mathbf{Y}_n$ is

$$M_n(\mathbf{t}) = E\left[\exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} W_i\right)\right],$$

we see that this is the moment generating function of $(1/\sqrt{n})\sum_{i=1}^{n} W_i$ evaluated at $t = 1$. So

$$M_n(\mathbf{t}) = E\left[\exp\left((1)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} W_i\right)\right] \to e^{(1)^2\mathbf{t}'\mathbf{\Sigma}\mathbf{t}/2} = e^{\mathbf{t}'\mathbf{\Sigma}\mathbf{t}/2}$$

by Theorem 5.4.3 and $(*)$, since the moment generating function of $N(0, \mathbf{t}'\mathbf{\Sigma}\mathbf{t})$ is $e^{t^2\mathbf{t}'\mathbf{\Sigma}\mathbf{t}/2}$ by Note 3.4.B. But $e^{\mathbf{t}'\mathbf{\Sigma}\mathbf{t}/2}$ is the moment generating function of a $N_p(\mathbf{0}, \mathbf{\Sigma})$ by Note 3.5.E, so by the uniqueness of the moment generating function (by Theorem 1.9.2) we have that $\mathbf{Y}_n \xrightarrow{D} N(\mathbf{0}, \mathbf{\Sigma})$, as claimed. □

Theorem 5.4.4 (continued 2)

Proof (continued). Since the moment generating function of $\mathbf{Y}_n$ is

$$M_n(\mathbf{t}) = E\left[\exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} W_i\right)\right],$$

we see that this is the moment generating function of $(1/\sqrt{n})\sum_{i=1}^{n} W_i$ evaluated at $t = 1$. So

$$M_n(\mathbf{t}) = E\left[\exp\left((1)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} W_i\right)\right] \to e^{(1)^2\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2} = e^{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2}$$

by Theorem 5.4.3 and $(*)$, since the moment generating function of $N(0, \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})$ is $e^{t^2\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2}$ by Note 3.4.B. But $e^{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2}$ is the moment generating function of a $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ by Note 3.5.E, so by the uniqueness of the moment generating function (by Theorem 1.9.2) we have that $\mathbf{Y}_n \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma})$, as claimed. □