6. Convergence of Random Variables

6.2 Types of convergence

X_n **converges to X in probability**, written $X_n \xrightarrow{\text{P}} X$, if, for every $\epsilon > 0$,:

$$\mathbb{P}(|X_n - X| > \epsilon) \to 0$$

as $n \to \infty$.

X_n **converges to X in distribution**, written $X_n \rightsquigarrow X$, if

$$\lim_{n \to \infty} F_n(t) = F(t)$$

for all t for which F is continuous.

X_n **converges to X in quadratic mean**, written $X_n \xrightarrow{\text{qm}} X$, if,

$$\mathbb{E}(X_n - X)^2 \to 0$$

as $n \to \infty$.

Theorem 6.4. The following relationships hold:

1. $X_n \xrightarrow{\text{qm}} X$ implies that $X_n \xrightarrow{P} X$.
2. $X_n \xrightarrow{\text{P}} X$ implies that $X_n \rightsquigarrow X$.
3. if $X_n \rightsquigarrow X$ and if $\mathbb{P}(X = c) = 1$ for some real number c, then $X_n \xrightarrow{\text{P}} X$.

Proof

1. Fix $\epsilon > 0$. Using Chebyshev's inequality,

$$\mathbb{P}(|X_n - X| > \epsilon) = \mathbb{P}(|X_n - X|^2 < \epsilon^2) \leq \frac{\mathbb{E}|X_n - X|^2}{\epsilon^2} \to 0$$

1. Fix $\epsilon > 0$ and let x be a point of continuity of F. Then

$$\begin{aligned} F_n(x) &= \mathbb{P}(X_n \leq x) = \mathbb{P}(X_n \leq x, X \leq x + \epsilon) + \mathbb{P}(X_n \leq x, X > x + \epsilon) && (1) \\ &\leq \mathbb{P}(X \leq x + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon) && (2) \\ &= F(x + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon) && (3) \end{aligned}$$

Also,

$$\begin{aligned} F(x - \epsilon) &= \mathbb{P}(X \leq x - \epsilon) = \mathbb{P}(X \leq x - \epsilon, X_n \leq x) + \mathbb{P}(X \leq x + \epsilon, X_n > x) && (4) \\ &\leq F_n(x) + \mathbb{P}(|X_n - X| > \epsilon) && (5) \end{aligned}$$

Hence,

$$F(x - \epsilon) - \mathbb{P}(|X_n - X| > \epsilon) \leq F_n(x) \leq F_n(x + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon)$$

Take the limit as $n \rightarrow \infty$ to conclude that

$$F(x - \epsilon) \leq \liminf_{n \rightarrow \infty} F_n(x) \leq \limsup_{n \rightarrow \infty} F_n(x) \leq F(x + \epsilon)$$

1. Fix $\epsilon > 0$. Then,

$$\begin{aligned} \mathbb{P}(|X_n - c| > \epsilon) &= \mathbb{P}(X_n < c - \epsilon) + \mathbb{P}(X_n > c + \epsilon) && (6) \\ &\leq \mathbb{P}(X_n \leq c - \epsilon) + \mathbb{P}(X_n > c + \epsilon) && (7) \\ &= F_n(c - \epsilon) + 1 - F_n(c + \epsilon) && (8) \\ &\rightarrow F(c - \epsilon) + 1 - F(c + \epsilon) && (9) \\ &= 0 + 1 - 1 = 0 && (10) \end{aligned}$$

Now, to show that the reverse implications do not hold:

Convergence in probability does not imply convergence in quadratic mean

Let $U \sim \text{Unif}(0, 1)$, and let $X_n \sim \sqrt{n} I_{(0, 1/n)}(U)$. Then $\mathbb{P}(|X_n| > \epsilon) = \mathbb{P}(\sqrt{n} I_{(0, 1/n)}(U) > \epsilon) = \mathbb{P}(0 \leq U < 1/n) = 1/n \rightarrow 0$. Hence, then $X_n \xrightarrow{\text{P}} 0$. But $\mathbb{E}(X_n^2) = n \int_0^{1/n} du = 1$ for all n so X_n does not converge in quadratic mean.

Convergence in distribution does not imply convergence in probability

Let $X \sim N(0,1)$. Let $X_n = -X$ for $n = 1, 2, 3, \ldots$; hence $X_n \sim N(0,1)$. X_n has the same distribution as X for all n so, trivially, $\lim_n F_n(x) \to F(x)$ for all x. Therefore, $X_n \rightsquigarrow X$. But $\mathbb{P}(|X_n - X| > \epsilon) = \mathbb{P}(|2X| > \epsilon) \neq 0$. So X_n does not tend to X in probability.

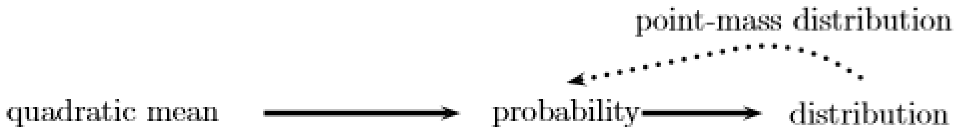

FIGURE 6.1. Relationship between types of convergence.

Theorem 6.5 Let X_n, X, Y_n, Y be random variables. Let g be a continuous function. Then:

1. If $X_n \overset{\text{P}}{\to} X$ and $Y_n \overset{\text{P}}{\to} Y$, then $X_n + Y_n \overset{\text{P}}{\to} X + Y$.
2. If $X_n \overset{\text{qm}}{\longrightarrow} X$ and $Y_n \overset{\text{qm}}{\longrightarrow} Y$, then $X_n + Y_n \overset{\text{qm}}{\longrightarrow} X + Y$.
3. If $X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow c$, then $X_n + Y_n \rightsquigarrow X + c$.
4. If $X_n \overset{\text{P}}{\to} X$ and $Y_n \overset{\text{P}}{\to} Y$, then $X_n Y_n \overset{\text{P}}{\to} XY$.
5. If $X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow c$, then $X_n Y_n \rightsquigarrow cX$.
6. If $X_n \overset{\text{P}}{\to} X$ then $g(X_n) \overset{\text{P}}{\to} g(X)$.
7. If $X_n \rightsquigarrow X$ then $g(X_n) \rightsquigarrow g(X)$.

6.3 The Law of Large Numbers

Theorem 6.6 (The Weak Law of Large Numbers (WLLN)). If $X_1, X_2, \ldots, X_n$ are IID, then $\overline{X}_n \overset{\text{P}}{\to} \mu$.

Proof: Assume that $\sigma < \infty$. This is not necessary but it simplifies the proof. Using Chebyshev's inequality,

$$\mathbb{P}(|\overline{X}_n - \mu| > \epsilon) \leq \frac{\mathbb{E}(|\overline{X}_n - \mu|^2)}{\epsilon^2} = \frac{\mathbb{V}(\overline{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

which tends to 0 as $n \to \infty$.

6.4 The Central Limit Theorem

Theorem 6.8 (The Central Limit Theorem (CLT)). Let $X_1, X_2, \ldots, X_n$ be IID with mean μ and variance σ^2. Let $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$. Then

$$Z_n \equiv \frac{\sqrt{n}\left(\overline{X}_n - \mu\right)}{\sigma} \rightsquigarrow Z$$

where $Z \sim N(0,1)$. In other words,

$$\lim_{n\to\infty} \mathbb{P}(Z_n \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

In addition to $Z_n \rightsquigarrow N(0,1)$, there are several forms of notation to denote the fact that the distribution of Z_n is converging to a Normal. They all mean the same thing. Here they are:

$$\begin{align} Z_n &\approx N(0,1) \tag{11} \\ \overline{X}_n &\approx N\left(\mu, \frac{\sigma^2}{n}\right) \tag{12} \\ \overline{X}_n - \mu &\approx N\left(0, \frac{\sigma^2}{n}\right) \tag{13} \\ \sqrt{n}(\overline{X}_n - \mu) &\approx N\left(0, \sigma^2\right) \tag{14} \\ \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} &\approx N(0,1) \tag{15} \end{align}$$

The central limit theorem tells us that $Z_n = \sqrt{n}(\overline{X}_n - \mu)/\sigma$ is approximately $N(0,1)$. However, we rarely know σ. We can estimate σ^2 from $X_1, X_2, \dots, X_n$ by

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

This raises the following question: if we replace σ with S_n is the central limit theorem still true? The answer is yes.

Theorem 6.10. Assume the same conditions as the CLT. Then,

$$\frac{\sqrt{n}\left(\overline{X}_n - \mu\right)}{S_n} \rightsquigarrow N(0,1)$$

You might wonder how accurate the normal approximation is. The answer is given by the Berry-Essèen theorem.

Theorem 6.11 (Berry-Essèen). Suppose that $\mathbb{E}|X_1|^3 < \infty$. Then

$$\sup_z |\mathbb{P}(Z_n \le z) - \Phi(z)| \le \frac{33}{4} \frac{\mathbb{E}|X_1 - \mu|^3}{\sqrt{n}\sigma^3}$$

There is also a multivariate version of the central limit theorem.

Theorem 6.12 (Multivariate central limit theorem). Let $X_1, \dots, X_n$ be IID random vectors where

$$X_i = \begin{pmatrix} X_{1i} \\ X_{2i} \\ \vdots \\ X_{ki} \end{pmatrix}$$

with mean

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} \mathbb{E}(X_{1i}) \\ \mathbb{E}(X_{2i}) \\ \vdots \\ \mathbb{E}(X_{ki}) \end{pmatrix}$$

and variance matrix Σ. Let

$$\overline{X} = \begin{pmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_k \end{pmatrix}$$

where $\overline{X}_r = n^{-1} \sum_{i=1}^n X_{ri}$. Then,

$$\sqrt{n}(\overline{X} - \mu) \rightsquigarrow N(0, \Sigma)$$

6.5 The Delta Method

Theorem 6.13 (The Delta Method). Suppose that

$$\frac{\sqrt{n}(Y_n - \mu)}{\sigma} \rightsquigarrow N(0,1)$$

and that g is a differentiable function such that $g'(u) \neq 0$. Then

$$\frac{\sqrt{n}(g(Y_n) - g(u))}{|g'(u)|\sigma} \rightsquigarrow N(0,1)$$

In other words,

$$Y_n \approx N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow g(Y_n) \approx N\left(g(\mu), (g'(\mu))^2 \frac{\sigma^2}{n}\right)$$

Theorem 6.15 (The Multivariate Delta method). Suppose that $Y_n = (Y_{n1}, \dots, Y_{nk})$ is a sequence of random vectors such that

$$\sqrt{n}(Y_n - \mu) \rightsquigarrow N(0, \Sigma)$$

Let $g : \mathbb{R}^k \to \mathbb{R}$ and let

$$\nabla g = \begin{pmatrix} \frac{\partial g}{\partial y1} \\ \vdots \\ \frac{\partial g}{\partial yk} \end{pmatrix}$$

Let ∇_μ denote $\nabla g(y)$ evaluated at $y = \mu$ and assume that the elements of ∇_μ are non-zero. Then

$$\sqrt{n}(g(Y_n) - g(\mu)) \rightsquigarrow N(0, \nabla_\mu^T \Sigma \nabla_\mu)$$

6.6 Technical appendix

X_n **converges to X almost surely**, written $X_n \xrightarrow{\text{as}} X$, if

$$\mathbb{P}(\{s : X_n(s) \to X(s)\}) = 1$$

X_n **converges to X in L_1**, written $X_n \xrightarrow{L1} X$, if

$$\mathbb{E}|X_n - X| \to 0$$

Theorem 6.17. Let X_n and X be random variables. Then:

1. $X_n \xrightarrow{\text{as}} X$ implies that $X_n \xrightarrow{\text{P}} X$.
2. $X_n \xrightarrow{\text{qm}} X$ implies that $X_n \xrightarrow{L_1} X$.
3. $X_n \xrightarrow{L_1} X$ implies that $X_n \xrightarrow{\text{P}} X$.

The weak law of large numbers says that $\overline{X}_n$ converges to $\mathbb{E}X$ in probability. The strong law asserts that this is also true almost surely.

Theorem 6.18 (The strong law of large numbers). Let $X_1, X_2, \dots, X_n$ be IID. If $\mu = \mathbb{E}|X_1| < \infty$ then $\overline{X}_n \xrightarrow{\text{as}} \mu$.

A sequence is **asymptotically uniformly integrable** if

$$\lim_{M \to \infty} \limsup_{n \to \infty} \mathbb{E}(|X_n| I(|X_n| > M)) = 0$$

If $X_n \xrightarrow{\text{P}} b$ and X_n is asymptotically uniformly integrable, then $\mathbb{E}(X_n) \to b$.

The **moment generating function** of a random variable X is

$$\psi_X(t) = \mathbb{E}(e^{tX}) = \int_u e^{tu} f_X(u) du$$

Lemma 6.19. Let $Z_1, Z_2, \dots, Z_n$ be a sequence of random variables. Let ψ_n be the mgf of Z_n. Let Z be another random variable and denote its mgf by ψ. If $\psi_n(t) \to \psi(t)$ for all t in some open interval around 0, then $Z_n \rightsquigarrow Z$.

Proof of the Central Limit Theorem

Let $Y_i = (X_i - \mu)/\sigma$. Then, $Z_n = n^{-1/2} \sum_i Y_i$. Let $\psi(t)$ be the mgf of Y_i. The mgf of $\sum_i Y_i$ is $(\psi(t))^n$ and the mgf of Z_n is $[\psi(t/\sqrt{n})]^n \equiv \xi_n(t)$.

Now $\psi'(0) = \mathbb{E}(Y_1) = 0$ and $\psi''(0) = \mathbb{E}(Y_1^2) = \mathbb{V}(Y_1) = 1$. So,

$$\begin{aligned}
\psi(t) &= \psi(0) + t\psi'(0) + \frac{t^2}{2!}\psi''(0) + \frac{t^3}{3!}\psi'''(0) + \dots & (16) \\
&= 1 + 0 + \frac{t^2}{2} + \frac{t^3}{3!}\psi'''(0) + \dots & (17) \\
&= 1 + \frac{t^2}{2} + \frac{t^3}{3!}\psi'''(0) + \dots & (18)
\end{aligned}$$

Now,

$$\xi_n(t) = \left[\psi\left(\frac{t}{\sqrt{n}}\right)\right]^n \tag{19}$$

$$= \left[1 + \frac{t^2}{2n} + \frac{t^3}{3!n^{3/2}}\psi'''(0) + \dots\right]^n \tag{20}$$

$$= \left[1 + \frac{\frac{t^2}{2} + \frac{t^3}{3!n^{1/2}}\psi'''(0) + \dots}{n}\right]^n \tag{21}$$

$$\to e^{t^2/2} \tag{22}$$

which is the mgf of $N(0,1)$. The resolt follows from the previous theorem. In the last step we used the fact that, if $a_n \to a$, then

$$\left(1 + \frac{a_n}{n}\right)^n \to e^a$$

6.8 Exercises

Exercise 6.8.1. Let $X_1, \dots, X_n$ be iid with finite mean $\mu = \mathbb{E}(X_i)$ and finite variance $\sigma^2 = \mathbb{V}(X_i)$. Let $\overline{X}_n$ be the sample mean and let S_n^2 be the sample variance.

(a) Show that $\mathbb{E}(S_n^2) = \sigma^2$.

Solution:

S_n^2 is the sample variance, that is, $S_n^2 = n^{-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2$. Therefore:

$$
\begin{aligned}
\mathbb{E}[S_n^2] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \left(X_i - \frac{1}{n}\sum_{j=1}^n X_j\right)^2\right] && (23)\\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[X_i^2 - \frac{2}{n}X_i\sum_{j=1}^n X_j + \frac{1}{n^2}\sum_{j=1}^n X_j \sum_{k=1}^n X_k\right] && (24)\\
&= \frac{1}{n}\sum_{i=1}^n \left[\frac{n-2}{n}\mathbb{E}[X_i^2] - \frac{2}{n}\sum_{j\neq i}\mathbb{E}[X_i X_j] + \frac{1}{n^2}\sum_{j=1}^n\sum_{k\neq j}\mathbb{E}[X_j X_k] + \frac{1}{n^2}\sum_{j=1}^n \mathbb{E}[X_j^2]\right] && (25)\\
&= \frac{1}{n}\sum_{i=1}^n \left[\frac{n-2}{n}(\sigma^2+\mu^2) - \frac{2}{n}(n-1)\mu^2 + \frac{1}{n^2}n(n-1)\mu^2 + \frac{1}{n}(\sigma^2+\mu^2)\right] && (26)\\
&= \frac{n-1}{n}\sigma^2 && (27)
\end{aligned}
$$

(b) Show that $S_n^2 \xrightarrow{\text{P}} \sigma^2$.

Hint: show that $S_n^2 = c_n n^{-1}\sum_{i=1}^n X_i^2 - d_n \overline{X}_n^2$ where $c_n \to 1$ and $d_n \to 1$. Apply the law of large numbers to $n^{-1}\sum_{i=1}^n X_i^2$ and to $\overline{X}_n$. Then use part (e) of Theorem 6.5.

Solution:

We have:

$$
\begin{aligned}
\overline{X}_n^2 &= \frac{1}{n^2}\left(\sum_{i=1}^n X_i\right)^2 && (28)\\
&= \frac{1}{n^2}\sum_{i=1}^n X_i^2 + \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1, i\neq j}^n X_i X_j && (29)
\end{aligned}
$$

Isolating the sum of products, we have:

$$
\sum_{j=1, i\neq j}^n X_i X_j = n^2 \overline{X}_n^2 - \sum_{i=1}^n X_i^2
$$

Now, from the sample variance :

$$\begin{aligned}
S_n^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 && (30)\\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2X_i\overline{X}_n + \overline{X}_n^2\right) && (31)\\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_j + \overline{X}_n^2 && (32)\\
&= \left(\frac{1}{n} - \frac{2}{n^2}\right)\sum_{i=1}^{n}X_i^2 - \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1, i\neq j}^{n}X_iX_j + \overline{X}_n^2 && (33)\\
&= \left(\frac{1}{n} - \frac{2}{n^2}\right)\sum_{i=1}^{n}X_i^2 - \frac{2}{n^2}\left(n^2\overline{X}_n^2 - \sum_{i=1}^{n}X_i^2\right) + \overline{X}_n^2 && (34)\\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \overline{X}_n^2 && (35)
\end{aligned}$$

We can use $c_n = 1$ and $d_n = 1$, as suggested in the hint. Applying the law of large numbers,

$$n^{-1}\sum_{i=1}^{n}X_i^2 \xrightarrow{\text{P}} n^{-1}\sum_{i=1}^{n}\mathbb{E}(X_i^2) = \sigma^2 + \mu^2$$

$$\overline{X}_n \xrightarrow{\text{P}} n^{-1}\sum_{i=1}^{n}\mathbb{E}(X_i) = \mu \Rightarrow \overline{X}_n^2 \xrightarrow{\text{P}} \mu^2$$

Therefore, from theorem 6.5.e, $S_n^2 = c_n n^{-1}\sum_{i=1}^{n}X_i^2 - d_n\overline{X}_n^2 \xrightarrow{\text{P}} \sigma^2 + \mu^2 - \mu^2 = \sigma^2$.

Exercise 6.8.2. Let $X_1, X_2, \dots, X_n$ be a sequence of random variables. Show that $X_n \xrightarrow{\text{qm}} b$ if and only if

$$\lim_{n\to\infty}\mathbb{E}(X_n) = b \quad \text{and} \quad \lim_{n\to\infty}\mathbb{V}(X_n) = 0 \tag{36}$$

Solution:

$X_n \xrightarrow{\text{qm}} b$ id equivalent to:

$$
\begin{aligned}
\mathbb{E}[(X_n - b)^2] \to 0 && (37)\\
\mathbb{E}[X_n^2 - 2bX_n + b^2] \to 0 && (38)\\
\mathbb{E}[X_n^2] - 2b\mathbb{E}[X_n] + b^2 \to 0 && (39)\\
\mathbb{E}[X_n^2] - 2b\mathbb{E}[X_n] + b^2 \to 0 && (40)\\
\mathbb{V}[X_n] + (\mathbb{E}[X_n])^2 - 2b\mathbb{E}[X_n] + b^2 \to 0 && (41)
\end{aligned}
$$

If $\lim_{n\to\infty} \mathbb{V}[X_n] = 0$ and $\lim_{n\to\infty} \mathbb{E}[X_n] = b$, then

$$
\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}[(X_n - b)^2] = && (42)\\
&= \lim_{n\to\infty} \mathbb{V}[X_n] + (\mathbb{E}[X_n])^2 - 2b\mathbb{E}[X_n] + b^2 && (43)\\
&= \lim_{n\to\infty} \mathbb{V}[X_n] + (\lim_{n\to\infty} \mathbb{E}[X_n])^2 - 2b \lim_{n\to\infty} \mathbb{E}[X_n] + b^2 && (44)\\
&= 0 + b^2 - 2b^2 + b^2 && (45)\\
&= 0 && (46)
\end{aligned}
$$

On the other direction, if $X_n \xrightarrow{\text{qm}} b$, then

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{V}[X_n] + (\lim_{n\to\infty} \mathbb{E}[X_n])^2 - 2b \lim_{n\to\infty} \mathbb{E}[X_n] + b^2 = 0 && (47)\\
\lim_{n\to\infty} \mathbb{V}[X_n] + (\lim_{n\to\infty} \mathbb{E}[X_n] - b)^2 = 0 && (48)\\
\lim_{n\to\infty} \mathbb{V}[X_n - b] + \lim_{n\to\infty} (\mathbb{E}[X_n - b])^2 = 0 && (49)
\end{aligned}
$$

Since both terms inside the limits are non-negative, the limits themselves are non-negative. Two non-negative values add up to 0, so they must both be zero, and so we have:

$$
\lim_{n\to\infty} \mathbb{E}(Y_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{V}(Y_n) = 0 \tag{50}
$$

or, equivalently,

$$
\lim_{n\to\infty} \mathbb{E}(X_n) = b \quad \text{and} \quad \lim_{n\to\infty} \mathbb{V}(X_n) = 0 \tag{51}
$$

Exercise 6.8.3. Let $X_1, X_2, \dots, X_n$ be iid and let $\mu = \mathbb{E}(X_i)$. Suppose that variance is finite. Show that $\overline{X}_n \xrightarrow{\text{qm}} \mu$.

Solution.

Let $Y_i = X_i - \mu$. It has variance $\sigma_Y = \sigma$ and mean $\mu_Y = 0$. We have:

$$
\begin{aligned}
&\mathbb{E}[(\overline{X}_n - \mu)^2] = && (52)\\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n (X_i - \mu)\right)^2\right] && (53)\\
&= \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^n Y_i\right)^2\right] && (54)\\
&= \frac{1}{n^2}\left(\sum_{i=1}^n \mathbb{E}[Y_i^2] - \sum_{i=1}^n \sum_{j=1, j\neq i}^n \mathbb{E}[Y_i Y_j]\right) && (55)\\
&= \frac{1}{n}\left((\sigma_Y^2 + \mu_Y^2) - (n-1)\mu_Y^2\right) && (56)\\
&= \frac{\sigma}{n} && (57)
\end{aligned}
$$

Therefore, $\lim_{n\to\infty} \mathbb{E}[(\overline{X}_n - \mu)^2] = \lim_{n\to\infty} \sigma/n = 0$, and so $\overline{X}_n \xrightarrow{\text{qm}} \mu$.

Exercise 6.8.4. Let $X_1, X_2, \dots$ be a sequence of random variables such that

$$
\mathbb{P}\left(X_n = \frac{1}{n}\right) = 1 - \frac{1}{n^2} \quad \text{and} \quad \mathbb{P}\left(X_n = n\right) = \frac{1}{n^2} \tag{58}
$$

Does X_n converge in probability? Doex X_n converge in quadratic mean?

Solution.

For any distribution X, we have:

$$
\begin{aligned}
&\mathbb{P}(|X_n - X| > \epsilon) = && (59)\\
&= \mathbb{P}\left(|X_n - X| > \epsilon \;\middle|\; X_n = \frac{1}{n}\right)\mathbb{P}\left(X_n = \frac{1}{n}\right) + \mathbb{P}\left(|X_n - X| > \epsilon \mid X_n = n\right)\mathbb{P}\left(X_n = n\right) && (60)\\
&= \mathbb{P}\left(\left|\frac{1}{n} - X\right| > \epsilon\right)\left(1 - \frac{1}{n^2}\right) + \mathbb{P}\left(|n - X| > \epsilon\right)\frac{1}{n^2} && (61)
\end{aligned}
$$

Looking at the limit as $n \to \infty$,

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = && (62)\\
&= \lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n} - X\right| > \epsilon\right)\left(1 - \frac{1}{n^2}\right) + \lim_{n\to\infty} \mathbb{P}\left(|n - X| > \epsilon\right)\frac{1}{n^2} && (63)\\
&= \lim_{n\to\infty} \mathbb{P}\left(|X| > \epsilon\right) && (64)
\end{aligned}$$

If we set $X = 0$, the limit above will be zero for any positive ϵ -- so we have $X_n \xrightarrow{\text{P}} 0$.

Now, for any quadratic mean potential convergence, we have:

$$\begin{aligned}
&\mathbb{E}\left[(X_n - X)^2\right] = && (65)\\
&= \mathbb{E}\left[(X_n - X)^2 \middle| X_n = \frac{1}{n}\right]\mathbb{P}\left(X_n = \frac{1}{n}\right) + \mathbb{E}\left[(X_n - X)^2 | X_n = n\right]\mathbb{P}\left(X_n = n\right) && (66)\\
&= \mathbb{E}\left[\left(X - \frac{1}{n}\right)^2\right]\left(1 - \frac{1}{n^2}\right) + \mathbb{E}\left[(X - n)^2\right]\frac{1}{n^2} && (67)\\
&= \mathbb{E}\left[X^2 - 2Xn^{-1} + n^{-2}\right]\left(1 - \frac{1}{n^2}\right) + \mathbb{E}\left[X^2 - 2Xn + n^2\right]\frac{1}{n^2} && (68)\\
&= \mathbb{E}\left[X^2\right] + \mathbb{E}\left[X\right]\left(\frac{-2}{n}\left(1 - \frac{1}{n^2}\right) - \frac{2}{n}\right) + \frac{1}{n^2}\left(1 - \frac{1}{n^2}\right) + 1 && (69)\\
&= \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]\frac{2}{n}\left(2 + \frac{1}{n^2}\right) + \frac{1}{n^2}\left(1 - \frac{1}{n^2}\right) + 1 && (70)
\end{aligned}$$

Taking the limit as $n \to \infty$,

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[(X_n - X)^2\right] = && (71)\\
&= 1 + \lim_{n\to\infty} \mathbb{E}\left[X^2\right] && (72)\\
&= 1 + \mathbb{E}\left[X^2\right] && (73)\\
&\geq 1 && (74)
\end{aligned}$$

so there is no distribution X for which this value is 0, and so there is no quadratic mean convergence.

Exercise 6.8.5. Let $X_1, \dots, X_n \sim \text{Bernoulli}(p)$. Prove that

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{\text{P}} p \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{\text{qm}} p \tag{75}$$

Solution.

Given that quadratic mean convergence implies probability convergence, we only need to prove the second proposition.

Let $Y_i = X_i^2 - p$. Then:

$$\begin{align}
\mathbb{E}[Y_i] &= \tag{76}\\
&= \mathbb{E}[X_i^2] - p \tag{77}\\
&= \mathbb{V}[X_i] + \mathbb{E}[X_i]^2 - p \tag{78}\\
&= p(1-p) + p^2 - p \tag{79}\\
&= 0 \tag{80}\\
\mathbb{E}[Y_i^2] &= \tag{81}\\
&= \mathbb{V}[Y_i] + \mathbb{E}[Y_i]^2 \tag{82}\\
&= \mathbb{V}[X_i^2 - p] + 0^2 \tag{83}\\
&= \mathbb{V}[X_i^2] + 0^2 \tag{84}\\
&= \mathbb{V}[X_i] \tag{85}\\
&= p(1-p) \tag{86}\\
\mathbb{E}[Y_i Y_j] &= \text{(for independent variables)} \tag{87}\\
&= \mathbb{E}[Y_i]\mathbb{E}[Y_j] \tag{88}\\
&= 0 \tag{89}
\end{align}$$

$$\mathbb{E}\left[\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - p\right)^2\right] = \tag{90}$$

$$= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - p\right)\right)^2\right] \tag{91}$$

$$= \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^2\right] \tag{92}$$

$$= \frac{1}{n^2}\mathbb{E}\left[\sum_{i=1}^{n} Y_i^2 - \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} Y_i Y_j\right] \tag{93}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathbb{E}\left[Y_i^2\right] - \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\mathbb{E}\left[Y_i Y_j\right]\right) \tag{94}$$

$$= \frac{p(1-p)}{n} \tag{95}$$

So, as $n \to \infty$, this expectation goes to 0, and we have quadratic mean convergence.

Exercise 6.8.6. Suppose that the height of men has mean 68 inches and standard deviation 4 inches. We draw 100 men at random. Find (approximately) the probability that the average height of men in our sample will be at least 68 inches.

Solution.

We assume all men's heights are measurements from iid variables X_i with mean $\mu = 68$ and variance $\sigma^2 = 16$.

We need to approximate $\mathbb{P}(\overline{X}_{100} > \mu)$. But by the central limit theorem,

$$\overline{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right)$$

so this probability will be approximately

$$\mathbb{P}\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \geq \frac{\sqrt{n}(\mu - \mu)}{\sigma}\right) = \mathbb{P}\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \geq 0\right) = P(Z \geq 0) = \frac{1}{2}$$

Exercise 6.8.7. Let $\lambda_n = 1/n$ for $n = 1, 2, \ldots$. Let $X_n \sim \text{Poisson}(\lambda_n)$.

(a) Show that $X_n \xrightarrow{\text{P}} 0$.

Solution.

$$\mathbb{E}(X_n^2) = \mathbb{V}(X_n) + \mathbb{E}(X_n)^2 = \lambda_n^2 + \lambda_n^2 = 2\lambda_n^2 = 2/n^2$$

This quantity goes to zero as $n \to \infty$, so we have $X_n \xrightarrow{\text{qm}} 0$, which implies $X_n \xrightarrow{\text{P}} 0$.

(b) Let $Y_n = nX_n$. Show that $Y_n \xrightarrow{\text{P}} 0$.

Solution.

$$\mathbb{E}(Y_n^2) = \mathbb{V}(Y_n) + \mathbb{E}(Y_n)^2 = n^2\lambda_n^2 + n^2\lambda_n^2 = 2n^2\lambda_n^2 = 2$$

so we *don't* have a straightforward quadratic mean convergence on Y_n.

We *don't* have a straightforward L_1 convergence either:

$$\mathbb{E}(|Y_n|) = \mathbb{E}(Y_n) = n\lambda_n = 1$$

However, we can show that $Y_n \rightsquigarrow 0$:

$$\lim_{n \to \infty} F_{Y_n}(t) = \lim_{n \to \infty} F_{Y_1}(t/n) = \lim_{n \to \infty} F_{Y_1}(t/n) = 0$$

as, when $n \to \infty$, the portion of the CDF in the positive neighborhood of 0 shrinks to $F_{Y_1}(0) = 0$.

We also have a point mass distribution on our target distribution $Y_\infty = 0$: probability of 1 in point 0, and 0 everywhere else.

Therefore, from theorem 6.4 item c, we have $Y_n \xrightarrow{\text{P}} 0$.

Exercise 6.8.8. Suppose we have a computer program consisting of $n = 100$ pages of code. Let X_i be the number of errors in the i-th page of code. Suppose that the X_i's are Poisson with mean 1 and that they are independent. Let $Y = \sum_{i=1}^n X_i$ be the total number of errors. Use the central limit theorem to approximate $\mathbb{P}(Y < 90)$.

Solution. We have $Y = n\overline{X}_n$, the total being n times the sample mean. We need to approximate:

We need to approximate $\mathbb{P}(\overline{X}_{100} < 0.9)$. But by the central limit theorem,

$$\overline{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right)$$

so this probability will be approximately

$$\mathbb{P}\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} < \frac{\sqrt{100}(0.9-1)}{0.1}\right) = \mathbb{P}\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} < -10\right) = P(Z < -10)$$

Exercise 6.8.9. Suppose that $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = 1/2$. Define

$$X_n = \begin{cases} X & \text{with probability } 1 - \frac{1}{n} \\ e^n & \text{with probability } \frac{1}{n} \end{cases} \tag{96}$$

Does X_n converge to X in probability? Does X_n converge to X in distribution? Does $\mathbb{E}(X - X_n)^2$ converge to 0?

Solution.

For any potential quadratic mean convergence, we'd have:

$$\begin{aligned}
&\mathbb{E}(X - X_n)^2 = & (97)\\
&= (\mathbb{E}(X - X_n | X_n = X)\mathbb{P}(X_n = X) + \mathbb{E}(X - X_n | X_n = e^n)\mathbb{P}(X_n = e^n))^2 & (98)\\
&= \left(\mathbb{E}(0)\left(1 - \frac{1}{n}\right) + \mathbb{E}(X - e^n)\frac{1}{n}\right)^2 & (99)\\
&= \frac{1}{n^2}\mathbb{E}(X - e^n)^2 & (100)\\
&= \frac{1}{n^2}(\mathbb{E}(X) - e^n)^2 & (101)\\
&= \frac{e^{2n}}{n^2} & (102)
\end{aligned}$$

which does not converge to 0, so we do not have quadratic mean convergence.

For any potential distribution convergence, X_n has a point mass distribution, and we can write its CDF F_{X_n} explicitly as:

$$F_{X_n}(t) = \begin{cases} 0 & \text{if } t < -1 \\ \frac{1}{2}\left(1 - \frac{1}{n}\right) & \text{if } -1 \leq t < 1 \\ 1 - \frac{1}{n} & \text{if } 1 \leq t < e^n \\ 1 & \text{if } e^n \leq t \end{cases} \tag{103}$$

On the other hand, the CDF F_X of the target distribution X is:

$$F_X(t) = \begin{cases} 0 & \text{if } t < -1 \\ \frac{1}{2} & \text{if } -1 \leq t < 1 \\ 1 & \text{if } 1 \leq t \end{cases} \tag{104}$$

We then have:

$$F_X(t) - F_{X_n}(t) = \begin{cases} 0 & \text{if } t < -1 \\ \frac{1}{2n} & \text{if } -1 \leq t < 1 \\ \frac{1}{n} & \text{if } 1 \leq t < e^n \\ 0 & \text{if } e^n \leq t \end{cases} \tag{105}$$

so $0 \leq F_X(t) - F_{X_n}(t) \leq 1/n$, which goes to 0 as $n \to \infty$. Therefore $\lim_{n \to \infty} F_{X_n}(t) = F_X(t)$, or $X_n \rightsquigarrow X$.

Distribution convergence implies probability convergence, so we also have probability convergence, $X_n \xrightarrow{\text{P}} X$.

Exercise 6.8.10. Let $Z \sim N(0, 1)$. Let $t > 0$.

(a) Show that, for any $k > 0$,

$$\mathbb{P}(|Z| > t) \leq \frac{\mathbb{E}|Z|^k}{t^k}$$

Solution.

We have:

$$\mathbb{E}|Z|^k = \tag{106}$$

$$= \int_{-\infty}^{\infty} |z|^{k+1} \left(\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right) dz \tag{107}$$

$$= \int_{-\infty}^{0} (-z)^{k+1} \left(\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right) dz + \int_{0}^{\infty} z^{k+1} \left(\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right) dz \tag{108}$$

$$= \left\{ \frac{2}{\pi} \right\}^{1/2} \int_{0}^{\infty} z^k \left(z e^{-z^2/2} \right) dz \tag{109}$$

For t > 0,

$$\mathbb{P}(|Z| > t) = \tag{110}$$

$$= 2 \int_{t}^{\infty} z \left(\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right) dz \tag{111}$$

$$= \left\{ \frac{2}{\pi} \right\}^{1/2} \int_{t}^{\infty} z e^{-z^2/2} dz \tag{112}$$

Now we need to prove:

$$\int_{t}^{\infty} z e^{-z^2/2} dz \leq \frac{1}{t^k} \int_{0}^{\infty} z^k \left(z e^{-z^2/2} \right) dz$$

As the integrands are always positive, we can prove the stronger statement that, for $k \geq 0$:

$$\int_{t}^{\infty} z e^{-z^2/2} dz \leq \frac{1}{t^k} \int_{t}^{\infty} z^k \left(z e^{-z^2/2} \right) dz \tag{113}$$

$$t^k \int_{t}^{\infty} z e^{-z^2/2} dz \leq \int_{t}^{\infty} z^k \left(z e^{-z^2/2} \right) dz \tag{114}$$

$$0 \leq \int_{t}^{\infty} (z^k - t^k) \left(z e^{-z^2/2} \right) dz \tag{115}$$

But that's true, since $(z^k - t^k)(z e^{-z^2/2}) \geq 0$ whenever $z \geq t$. So the given statement follows.

(b) (Mill's inequality) Show that

$$\mathbb{P}(|Z| > t) \leq \left\{ \frac{2}{\pi} \right\}^{1/2} \frac{e^{-t^2/2}}{t}$$

Hint. Note that $\mathbb{P}(|Z| > t) = 2\mathbb{P}(Z > t)$. Now write out what $\mathbb{P}(Z > t)$ means and note that $x/t > 1$ whenever $x > t$.

Solution.

The stronger result we proved in (a) was, for $k \geq 0$,

$$\mathbb{P}(|Z| > t) = \left\{ \frac{2}{\pi} \right\}^{1/2} \int_t^\infty z e^{-z^2/2} dz \leq \left\{ \frac{2}{\pi} \right\}^{1/2} \frac{1}{t^k} \int_t^\infty z^k \left(z e^{-z^2/2} \right) dz$$

If we use $k = 0$, we get:

$$\mathbb{P}(|Z| > t) \leq \left\{ \frac{2}{\pi} \right\}^{1/2} \frac{1}{t} \int_t^\infty z e^{-z^2/2} dz = \left\{ \frac{2}{\pi} \right\}^{1/2} \frac{e^{-t^2/2}}{t}$$

which is the desired result.

Exercise 6.8.11. Suppose that $X_n \sim N(0, 1/n)$ and let X be a random variable with distribution $F(x) = 0$ if $x < 0$ and $F(x) = 1$ if $x \geq 0$. Does X_n converge to X in probability? Does X_n converge to X in distribution?

Solution.

We do not have convergence in distribution: $F_{X_n}(0) = 1/2$ for any n (as the normal distribution is symmetric around its mean), so $\lim_{n \rightarrow \infty} F_{X_n}(0) = 1/2 \neq F_X(0) = 1$.

We do have convergence in probability: for every $\epsilon > 0$,

$$\mathbb{P}(|X - X_n| > \epsilon) = \mathbb{P}(|X_n| > \epsilon) = 2\mathbb{P}(X_n > \epsilon) = 2(1 - F_{X_n}(\epsilon))$$

so

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X - X_n| > \epsilon) = 2(1 - \lim_{n \rightarrow \infty} F_{X_n}(\epsilon)) = 2(1 - \lim_{n \rightarrow \infty} F_{X_1}(n\epsilon)) = 2(1 - 1) = 0$$

Exercise 6.8.12. Let $X, X_1, X_2, X_3, \cdots$ be random variables that are positive and integer valued. Show that $X_n \rightsquigarrow X$ if and only if

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$$

for every integer k.

Solution.

If $X_n \rightsquigarrow X$, then $\lim_{n \rightarrow \infty} F_{X_n}(k) = F_X(k)$ for every integer k. But since the variables are positive and integer valued,

$$\mathbb{P}(X_n = k) = F_{X_n}(k) - F_{X_n}(k-1) \quad \text{and} \quad \mathbb{P}(X = k) = F_X(k) - F_X(k-1) \tag{116}$$

Therefore,

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \lim_{n \rightarrow \infty} F_{X_n}(k) - F_{X_n}(k-1) = \lim_{n \rightarrow \infty} F_{X_n}(k) - \lim_{n \rightarrow \infty} F_{X_n}(k-1) = F_X(k) - F_X(k-1) = \mathbb{P}(X = k)$$

On the other direction, if $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$, then

$$\lim_{n \rightarrow \infty} (F_{X_n}(k) - F_{X_n}(k-1)) = F_X(k) - F_X(k-1)$$

But the variables are positive and integer valued, so $F_{X_n}(k) = F_X(k) = 0$ for $k \leq 0$. We can then show that $\lim_{n \rightarrow \infty} F_{X_n}(k) = F_X(k)$ for every integer valued k by induction in k:

$$\lim_{n \rightarrow \infty} (F_{X_n}(k) - F_{X_n}(k-1)) = \left(\lim_{n \rightarrow \infty} F_{X_n}(k)\right) - F_X(k-1) = F_X(k) - F_X(k-1)$$

$$\Rightarrow \lim_{n \rightarrow \infty} F_{X_n}(k) = F_X(k)$$

Since the result holds for every integer variable k and the random variables can only take integer values, it must hold for all values, therefore $X_n \rightsquigarrow X$.

Exercise 6.8.13. Let $Z_1, Z_2, \dots$ be iid random variables with density f. Suppose that $\mathbb{P}(Z_i > 0) = 1$ and that $\lambda = \lim_{x \downarrow 0} f(x) > 0$. Let

$$X_n = n \min\{Z_1, \dots, Z_n\}$$

Show that $X_n \rightsquigarrow Z$ where Z has and exponential distribution with mean $1/\lambda$.

Solution.

Since $\mathbb{P}(Z_i > 0) = 1$, the cumulative density functions F assume value 0 for values up until 0 inclusive.

We have:

$$\mathbb{P}(X_n > x) = \mathbb{P}(n \min\{Z_1, \dots, Z_n\} > x) = \prod_{i=1}^n \mathbb{P}(Z_i > x/n)$$

Expanding the probability based on its density function,

$$\mathbb{P}(X_n > x) = \prod_{i=1}^n \mathbb{P}(Z_i > x/n) = \prod_{i=1}^n \int_0^{x/n} f(u)du = \left(F\left(\frac{x}{n}\right)\right)^n = \left(F(0) + F'(0)\frac{x}{n} + F''(0)\left(\frac{x}{n}\right)^2 \frac{1}{2!} + \cdots\right)^n$$

Taking the limit as $n \to \infty$,

$$\lim_{n \to \infty} \mathbb{P}(X_n > x) = \lim_{n \to \infty} \left(F(0) + F'(0)\frac{x}{n} + F''(0)\left(\frac{x}{n}\right)^2 \frac{1}{2!} + \cdots\right)^n = \lim_{n \to \infty} \left(F(0) + F'(0)\frac{x}{n}\right)^n = \lim_{n \to \infty} \left(0 + \lambda\frac{x}{n}\right)^n = e^{-\lambda x}$$

On the other hand, $\mathbb{P}(Z > x) = e^{-\lambda x}$, so the limit of the CDF complements are the same, and so $X_n \rightsquigarrow Z$.

Exercise 6.8.14. Let $X_1, \dots, X_n \sim \text{Uniform}(0, 1)$. Let $Y_n = \overline{X}_n^2$. Find the limiting distribution of Y_n.

Solution.

Let $F_K(x)$ denote the CDF of random variable K.

The sample mean $\overline{X}_n$ has a limiting distribution of $X = 1/2$, by the (strong) law of large numbers.

Then, for $x \geq 0$,

$$\mathbb{P}(Y_n > x) = \mathbb{P}(\overline{X}_n^2 > x) = \mathbb{P}(\overline{X}_n > x^{1/2})$$

$$F_{Y_n}(x) = F_{\overline{X}_n}(x^{1/2})$$

Since $\overline{X}_n \rightsquigarrow 1/2$,

$$\lim_{n \to \infty} F_{\overline{X}_n}(x) = F_{1/2}(x) \tag{117}$$
$$\lim_{n \to \infty} F_{\overline{X}_n}(x^{1/2}) = F_{1/2}(x^{1/2}) \tag{118}$$
$$\lim_{n \to \infty} F_Y(x) = F_{1/2}(x^{1/2}) \tag{119}$$

Therefore, $Y \rightsquigarrow Z$, where $F_Z(x) = F_{1/2}(x^{1/2})$, that is,

$$F_Z(t) = \begin{cases} 0 & \text{if } t^{1/2} < 1/2 \\ 1 & \text{otherwise} \end{cases} = \begin{cases} 0 & \text{if } t < 1/4 \\ 1 & \text{otherwise} \end{cases} \tag{120}$$

so Z assumes the constant value of $1/4$, and $Y \rightsquigarrow 1/4$.

Exercise 6.8.15. Let

$$\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}, \begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix}, \dots, \begin{pmatrix} X_{1n} \\ X_{1n} \end{pmatrix}$$

be iid random vectors with mean $\mu = (\mu_1, \mu_2)$ and variance Σ.

Let

$$\overline{X}_1 = \frac{1}{n} \sum_{i=1}^n X_{1i}, \quad \overline{X}_2 = \frac{1}{n} \sum_{i=1}^n X_{2i}$$

and define $Y_n = \overline{X}_1 / \overline{X}_2$. Find the limiting distribution of Y_n.

Solution.

Let

$$\overline{X} = n^{-1} \sum_{i=1}^n \begin{pmatrix} X_{1i} \\ X_{2i} \end{pmatrix} = \begin{pmatrix} \overline{X}_1 \\ \overline{X}_2 \end{pmatrix}$$

.

By the multivariate central limit theorem,

$$\sqrt{n}(\overline{X} - \mu) \rightsquigarrow N(0, \Sigma)$$

Define $g : \mathbb{R} \times \mathbb{R}_{\neq 0} \to \mathbb{R}$ as:

$$g \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_1 / y_2$$

Then, $Y_n = g(\overline{X_n})$ in every scenario where Y_n is defined.

Applying the multivariate delta method,

$$\nabla g = \begin{pmatrix} \frac{\partial g}{\partial y1} \\ \frac{\partial g}{\partial y2} \end{pmatrix} = \begin{pmatrix} \frac{1}{y2} \\ -\frac{y1}{y_2^2} \end{pmatrix}$$

Then,

$$\nabla_\mu = \begin{pmatrix} \frac{1}{\mu 2} \\ -\frac{\mu 1}{\mu_2^2} \end{pmatrix}$$

and so

$$\sqrt{n}(Y_n - g(\mu)) \rightsquigarrow N(0, \nabla_\mu^T \Sigma \nabla_\mu) \tag{121}$$
$$Y_n \rightsquigarrow N(\mu_1/\mu_2, n^{-1/2} \nabla_\mu^T \Sigma \nabla_\mu) \tag{122}$$