Chapter 5 Solutions

Andrew Wu
Wasserman: All of Statistics

February 24, 2025

Problem 5.1. Let $X_1, \dots, X_n$ be IID with finite mean $\mu = \mathbb{E}(X_1)$ and finite variance $\sigma^2 = \mathbb{V}(X_1)$. Let $\overline{X}_n$ be the sample mean and let S_n^2 be the sample variance.

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

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

Solution. For part (a), refer to the solution to problem 3.8 of this textbook.

For part (b): we have

$$\begin{aligned} S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i^2 - 2X_i\overline{X}_n + \overline{X}_n^2\right) \\ &= \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{2\overline{X}_n}{n-1}\sum_{i-1}^{n} X_i + \frac{n\overline{X}_n^2}{n-1} \\ &= \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{n\overline{X}_n^2}{n-1}. \end{aligned}$$

By the Law of Large Numbers, we know that $\sum_{i=1}^n X_i^2 \xrightarrow{P} n\mathbb{E}(X_1^2)$ and $\overline{X}_n \xrightarrow{P} \mathbb{E}(X_1)$. So as $n \to \infty$, $S_n^2 \xrightarrow{P} \frac{n}{n-1}\mathbb{V}(X_1) = \frac{n\sigma^2}{n-1}$. But as $n \to \infty$, $\frac{n}{n-1} \to 1$, so by Slutsky's theorem $S_n^2 \xrightarrow{P} \sigma^2$, as desired. $\square$

Problem 5.2. Let $X_1, X_2, \dots$ be a sequence of random variables. Show that $X_n \xrightarrow{\text{qm}} b$ if and only if

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

Solution. Assume firstly that $\lim_{n\to\infty}\mathbb{E}(X_n) = b$ and $\lim_{n\to\infty}\mathbb{V}(X_n) = 0$. The latter statement implies $\mathbb{E}(X_n^2) - \mathbb{E}(X_n)^2 \to 0$, so it follows that $\mathbb{E}(X_n^2) \to b^2$. Therefore it follows that $\mathbb{E}(X_n^2) - 2b\mathbb{E}(X_n) + b^2 \to b^2 - 2b^2 + b^2 = 0$, and the left-hand-side is equivalent to $\mathbb{E}[(X_n - b)^2]$ via linearity of expectation. Thus $X_n \xrightarrow{\text{qm}} b$.

Now assume $X_n \xrightarrow{\text{qm}} b$. Then $\mathbb{E}[(X_n - b)^2] \to 0$. But we have

$$\begin{aligned} \mathbb{E}[(X_n - b)^2] &= \mathbb{E}(X_n^2) - 2b\mathbb{E}(X_n) + b^2 \\ &= \mathbb{E}(X_n^2) - \mathbb{E}(X_n)^2 + \mathbb{E}(X_n)^2 - 2b\mathbb{E}(X_n) + b^2 \\ &= \mathbb{V}(X_n) + (\mathbb{E}(X_n) - b)^2. \end{aligned}$$

Thus $\mathbb{V}(X_n) + (\mathbb{E}(X_n) - b)^2 \to 0$, and as we know that $\mathbb{V}(X) \geq 0$ for any random variable X and that squares are nonnegative, it follows that $\mathbb{V}(X_n) \to 0$ and that $\mathbb{E}(X_n) \to b$. $\square$

Problem 5.3. Let $X_1, \ldots, X_n$ be IID and let $\mu = \mathbb{E}(X_1)$. Suppose that the variance is finite. Show that $\overline{X}_n \xrightarrow{\text{qm}} \mu$.

Solution. We will use the previous problem.

We have $\mathbb{E}(\overline{X}_n) = \mathbb{E}[\frac{1}{n}(X_1 + \ldots + X_n)] = \frac{1}{n} \cdot n \cdot \mu = \mu$, by linearity of expectation. Thus $\lim_{n\to\infty} \mathbb{E}(\overline{X}_n) = \mu$.

Moreover, we have $\mathbb{V}(\overline{X}_n) = \mathbb{V}[\frac{1}{n}(X_1 + \ldots + X_n)] = \frac{1}{n^2} \cdot n \cdot \mathbb{V}(X_1) = \frac{1}{n}\mathbb{V}(X_1)$. Then, knowing the variance is finite, we have $\lim_{n\to\infty} \mathbb{V}(\overline{X}_n) = \lim_{n\to\infty} \frac{1}{n}\mathbb{V}(X_1) = 0$.

It follows that as $\lim_{n\to\infty} \mathbb{E}(\overline{X}_n) \to \mu$ and $\lim_{n\to\infty} \mathbb{V}(\overline{X}_n) \to 0$ that $\overline{X}_n \xrightarrow{\text{qm}} \mu$. □

Problem 5.4. Let $X_1, X_2, \ldots$ be a sequence of random variables such that

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

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

Solution. We claim that X_n converges in probability, to 0, but that X_n does not converge in quadratic mean.

Note that $\mathbb{P}(|X_n - 0| > \epsilon) = \mathbb{P}(X_n = n) = \frac{1}{n^2}$ for all sufficiently large values of n such that $\frac{1}{n} < \epsilon$, so as $n \to \infty$, $\mathbb{P}(|X_n|) \to 0$. Thus $X_n \xrightarrow{P} 0$.

However,

$$\begin{aligned}
\mathbb{V}(X_n) &= \mathbb{E}(X_n^2) - \mathbb{E}(X_n)^2 \\
&= \left[\left(1 - \frac{1}{n^2}\right) \cdot \frac{1}{n^2} + \left(\frac{1}{n^2}\right) \cdot n^2\right] - \left[\left(1 - \frac{1}{n^2}\right) \cdot \frac{1}{n} + \left(\frac{1}{n^2}\right) \cdot n\right]^2 \\
&= \left[1 + \frac{1}{n^2} - \frac{1}{n^4}\right] - \frac{1}{n^2} = 1 - \frac{1}{n^4}
\end{aligned}$$

so $\lim_{n\to\infty} \mathbb{V}(X_n) \to 1$, so X_n does not converge in quadratic mean. □

Problem 5.5. Let $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$. Prove that

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

Solution. Given that $X_i \sim \text{Bernoulli}(p)$, we know that $X_i = 0$ or $X_i = 1$, so $X_i^2 = X_i$. It follows that $\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \overline{X}_n \xrightarrow{P} \mathbb{E}(X_i) = p$, by the Law of Large Numbers.

To show that $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{\text{qm}} p$, we'll show that $\overline{X}_n \xrightarrow{\text{qm}} p$, or equivalently that $\lim_{n\to\infty} \mathbb{E}(\overline{X}_n) = p$ and $\lim_{n\to\infty} \mathbb{V}(\overline{X}_n) = 0$.

We have $\mathbb{E}(\overline{X}_n) = \frac{1}{n} \cdot n \cdot \mathbb{E}(X_i) = p$, so $\lim_{n\to\infty} \mathbb{E}(\overline{X}_n) \to p$, as desired. Next, we have

$$\begin{aligned}
\mathbb{V}(\overline{X}_n) &= \frac{1}{n^2}\mathbb{V}(X_1 + \ldots + X_n) \\
&= \frac{1}{n}\mathbb{V}(X_i) \\
&= \frac{1}{n}(p - p^2)
\end{aligned}$$

so thus $\lim_{n\to\infty} \mathbb{V}(\overline{X}_n) \to 0$, too. It follows indeed that $\overline{X}_n \xrightarrow{\text{qm}} p$, and thus that $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{\text{qm}} p$.

(Note that we could've just proven $\overline{X}_n \xrightarrow{\text{qm}} p$, as it would then follow that $\overline{X}_n \xrightarrow{P} p$.) □

Problem 5.6. Suppose that the height of men has mean 68 inches and standard deviation 2.6 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. From now on, we denote X_n converges to Y in distribution by $X_n \xrightarrow{\mathrm{d}} Y$.

Denote the heights of the men by $X_1, \ldots, X_{100}$; note that their heights are all drawn from a distribution with mean 68 inches and variance 2.6^2 inches.

We know that $\overline{X}_n \xrightarrow{\mathrm{d}} N\left(\mu, \frac{\sigma^2}{n}\right)$ as $n \to \infty$, where in this case $\mu = 68$ and $\frac{\sigma^2}{n} = \frac{2.6^2}{100}$. As this is a normal distribution, it is symmetric around μ, so $\mathbb{P}(\overline{X}_n \geq 68) = 0.5$. □

Problem 5.7. Let $\lambda_n = 1/n$ for $n = 1, 2, \ldots$. Let $X_n \sim \text{Pois}(\lambda_n)$.

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

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

Solution. For part (a), we want to show that $\mathbb{P}(|X_n| > \epsilon) \to 0$ as $n \to \infty$, for any $\epsilon > 0$.

We have $\mathbb{P}(|X_n| > \epsilon) \leq \mathbb{P}(X_n \geq 1)$ for any $\epsilon > 0$. Let f_n be the PMF of the Poisson distribution with parameter $\frac{1}{n}$; then, $\mathbb{P}(X_n \geq 1) = \sum_{k=1}^{\infty} e^{-\frac{1}{n}} \frac{\left(\frac{1}{n}\right)^k}{k!}$. But we have

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=1}^{\infty} e^{-\frac{1}{n}} \frac{\left(\frac{1}{n}\right)^k}{k!} &= \lim_{n\to\infty} \left(\left[\sum_{k=0}^{\infty} e^{-\frac{1}{n}} \frac{\left(\frac{1}{n}\right)^k}{k!} \right] - e^{-\frac{1}{n}} \right) \\
&= \lim_{n\to\infty} (1 - e^{-\frac{1}{n}}) \\
&= 0.
\end{aligned}$$

Thus $\lim_{n\to\infty} \mathbb{P}(|X_n| > \epsilon) = 0$.

For part (b), we want to show that $\mathbb{P}(|nX_n| > \epsilon) \to 0$ as $n \to \infty$, for any $\epsilon > 0$.

We have $\mathbb{P}(|nX_n| > \epsilon) \leq \mathbb{P}(nX_n \geq 1) = \mathbb{P}(X_n \geq \frac{1}{n}) = \mathbb{P}(X_n \geq 1)$ for any $\epsilon > 0$. Thus $\lim_{n\to\infty} \mathbb{P}(|nX_n| > \epsilon) = 0$, too, by our previous work on part (a), and we are done. □

Problem 5.8. Suppose we have a computer program consisting of $n = 100$ pages of code. Let X_i be the number of errors on the ith page of code. Suppose that the X_is 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. The Central Limit Theorem states that $\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \xrightarrow{\mathrm{d}} Z$, so taking $n = 100$, $\mu = 1$, and $\sigma = 1$, we see that

$$10(\overline{X}_n - 1) \xrightarrow{\mathrm{d}} Z.$$

Note that $\overline{X}_n = \frac{Y}{n} = \frac{Y}{100}$, so it follows that

$$\frac{Y}{10} - 10 \xrightarrow{\mathrm{d}} Z.$$

Now we want to find $\mathbb{P}(Y < 90)$, or equivalently, $\mathbb{P}(Z < -1) = F_Z(-1)$. A quick check with the calculator reveals $F_Z(-1) = 0.159$. □

Problem 5.9. Suppose that $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = \frac{1}{2}$. Define

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

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

Solution. We claim that X_n converges to X in probability and distribution, but not in quadratic mean.

We have to show that $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$. But $|X_n - X| > \epsilon$ can occur only when $X_n = e^n$, so $\mathbb{P}(|X_n - X| > \epsilon) \geq \mathbb{P}(X_n = e^n) = \frac{1}{n}$, which approaches 0 as $n \to \infty$. Thus X_n converges to X in probability. Thus X_n must also converge to X in distribution.

Now we compute $\mathbb{E}((X-X_n)^2)$. Note that

$$(X-X_n)^2 = \begin{cases} 0, & \text{with probability } 1-\frac{1}{n} \\ (1-e^n)^2, & \text{with probability } \frac{1}{2n} \\ (-1-e^n)^2, & \text{with probability } \frac{1}{2n}. \end{cases}$$

Thus $\mathbb{E}((X-X_n)^2) = \frac{1}{n} + \frac{e^{2n}}{n}$, which does not converge to 0 as $n \to \infty$. □

Problem 5.10. Let $Z \sim N(0,1)$. Let $t > 0$. Show that for any $k > 0$,

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

Then compare this to Mill's inequality.

Solution. By Markov's inequality, we have $\mathbb{P}(|Z|^k > t^k) \leq \frac{\mathbb{E}(|Z|^k)}{t^k}$, but as $|Z|$ and t are both positive we have $\mathbb{P}(|Z| > t) = \mathbb{P}(|Z|^k > t^k)$, and we are done.

Comparing to the Mill bound, note that for $k = 1$ this bound is just the Markov bound and is quite weak. When k is large, this bound is also quite weak, as $\mathbb{E}(|Z|^k) \sim k^{k/2}$. This bound could be situationally better than the Mill bound for small values of k, for certain values of t. □

Problem 5.11. 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? In distribution?

Solution. Our random variable X clearly takes on the value 0 with probability 1.

We claim that $X_n \xrightarrow{P} 0$. We want to show that $\mathbb{P}(|X_n| > \epsilon) \to 0$. Note that by Markov's inequality, $\mathbb{P}(|X_n| > \epsilon) = \mathbb{P}(|X_n|^2 > \epsilon^2) \leq \frac{\mathbb{E}(|X_n|^2)}{\epsilon^2}$. But $\mathbb{E}(|X_n|^2) = \mathbb{E}(X_n^2)$, and $\mathbb{E}(X_n^2) - \mathbb{E}(X_n)^2 = \mathbb{V}(X_n) = \frac{1}{n}$, and thus as $\mathbb{E}(X_n) = 0$, it follows that $\mathbb{E}(|X_n|^2) = \frac{1}{n}$.

Therefore $\mathbb{P}(|X_n| > \epsilon) \leq \frac{1}{n\epsilon^2}$, and evidently as $n \to \infty$ we have $\frac{1}{n\epsilon^2} \to 0$, as desired. Moreover, this then implies that $X_n \xrightarrow{\mathrm{d}} 0$, too. □

Problem 5.12. Let $X, X_1, X_2, X_3, \ldots$ be random variables that are positive and integer-valued. Show that X_n converges in distribution to X if and only if

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

for every integer k.

Solution. Suppose firstly that $X_n \xrightarrow{\mathrm{d}} X$, and as usual, let F_n be the CDF of the random variable X_n. Then $\lim_{n\to\infty} F_n(t) = F(t)$ at all t for which F is continuous.

We proceed by strong induction. As $X, X_1, X_2, \ldots$ are positive and integer-valued, it follows that F is continuous at all non-integer positive reals. We see that

$$\begin{aligned} \lim_{n\to\infty} \mathbb{P}(X_n = 1) &= \lim_{n\to\infty} F_n(1.5) \\ &= F(1.5) \\ &= \mathbb{P}(X = 1) \end{aligned}$$

using the fact that $X, X_1, X_2, \ldots$ take on only positive integer values.

Now, suppose that for $k = 1, 2, \ldots, m$ that $\lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$. We claim that $\lim_{n\to\infty} \mathbb{P}(X_n = m+1) = \mathbb{P}(X = m+1)$.

Observe that $\lim_{n\to\infty} F_n(m+1.5) = F(m+1.5)$. But we also have

$$\begin{aligned} \lim_{n\to\infty} F_n(m+1.5) &= \lim_{n\to\infty} [\mathbb{P}(X_n = 1) + \mathbb{P}(X_n = 2) + \ldots + \mathbb{P}(X_n = m+1)] \\ &= \sum_{i=1}^{m+1} \lim_{n\to\infty} \mathbb{P}(X_n = i) \\ &= \left(\sum_{i=1}^{m} \mathbb{P}(X = i) \right) + \lim_{n\to\infty} \mathbb{P}(X_n = m+1). \end{aligned}$$

by our inductive hypothesis.

But

$$F(m+1.5) = \sum_{i=1}^{m+1} \mathbb{P}(X = i)$$

and thus it follows that $\lim_{n\to\infty} \mathbb{P}(X_n = m+1) = \mathbb{P}(X = m+1)$.

Now we show the opposite direction. Assume that $\lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$ for every integer k. We claim then that $\lim_{n\to\infty} F_n(t) = F(t)$ at all t for which F is continuous.

Evidently $F(t) = F_n(t) = 0$ for all positive integers n and $t < 1$, as we know that the random variables are positive integer-valued.

Let t be a positive real number at which F is not continuous; that is, either t is non-integer, or t is an integer for which $f(t) = 0$. We will show that $\lim_{n\to\infty} F_n(t) = F(t)$. Let $u = t-1$ if t is an integer; otherwise let $u = \lfloor t \rfloor$. Note that $F(t) = \mathbb{P}(X=1) + \mathbb{P}(X=2) + \ldots + \mathbb{P}(X=u)$, and that

$$\begin{aligned}
\lim_{n\to\infty} F_n(t) &= \lim_{n\to\infty} [\mathbb{P}(X_n = 1) + \mathbb{P}(X_n = 2) + \ldots + \mathbb{P}(X_n = u)] \\
&= \sum_{i=1}^{u} \lim_{n\to\infty} \mathbb{P}(X_n = i) \\
&= \sum_{i=1}^{u} \mathbb{P}(X = i)
\end{aligned}$$

so $\lim_{n\to\infty} F_n(t) = F(t)$, as desired. We are done. □

Problem 5.13. Let $Z_1, Z_2, \ldots$ 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

$$X_n = n \min\{Z_1, \ldots, Z_n\}.$$

Show that $X_n \xrightarrow{\text{d}} Z$, where Z has an exponential distribution with mean $1/\lambda$.

Solution. Let F be the CDF of $Z_1, Z_2, \ldots$, and let F_Z be the CDF of Z. Let F_n denote the CDF of X_n.

We want to show that $\lim_{n\to\infty} F_n(t) = F_Z(t)$. Note that

$$\begin{aligned}
F_n(t) &= \mathbb{P}(X_n < t) \\
&= \mathbb{P}(n \min(Z_1, \ldots, Z_n) < t) \\
&= \mathbb{P}\left(\min(Z_1, \ldots, Z_n) < \frac{t}{n}\right) \\
&= 1 - \prod_{i=1}^{n} \mathbb{P}\left(Z_i > \frac{t}{n}\right) \\
&= 1 - \left(1 - F\left(\frac{t}{n}\right)\right)^n.
\end{aligned}$$

Next, note that when $n \to \infty$, $\frac{t}{n} \to 0$, and thus

$$\lim_{n\to\infty} F\left(\frac{t}{n}\right) = \lim_{n\to\infty} \int_0^{\frac{t}{n}} f(x)dx = \frac{\lambda t}{n}$$

where we use that $\lambda = \lim_{x\downarrow 0} f(x)$.

Thus it follows that

$$\begin{aligned}
\lim_{n\to\infty} F_n(t) &= \lim_{n\to\infty} 1 - \left(1 - F\left(\frac{t}{n}\right)\right)^n \\
&= \lim_{n\to\infty} 1 - \left(1 - \frac{\lambda t}{n}\right)^n \\
&= 1 - e^{-\lambda t} = F_Z(t)
\end{aligned}$$

as desired. □

Problem 5.14. Let $X_1, \ldots, X_n \sim \text{Uniform}(0,1)$. Let $Y_n = \overline{X}_n^2$. Find the limiting distribution of Y_n.

Solution. We know that for large n, $\overline{X}_n \sim N(\mu, \sigma^2/n)$, where μ and σ^2 are the mean and variance of $\text{Uniform}(0,1)$, by the Central Limit Theorem.

Then, letting $g(x) = x^2$, we have by the Delta Method that when n is large,

$$g(\overline{X}_n) \sim N\left(g(\mu), (g'(\mu))^2 \cdot \frac{\sigma^2}{n}\right).$$

Substituting $\mu = \frac{1}{2}, \sigma^2 = \frac{1}{12}$, we obtain

$$Y_n \sim N\left(\frac{1}{4}, \frac{1}{12n}\right)$$

as the limiting distribution of Y_n. $\square$

Problem 5.15. Let

$$\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}, \begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix}, \ldots, \begin{pmatrix} X_{1n} \\ X_{2n} \end{pmatrix}$$

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

$$X_1 = \frac{1}{n}\sum_{i=1}^{n} X_{1i}, \quad \overline{X}_2 = \frac{1}{n}\sum_{i=1}^{n} X_{2i}$$

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

Solution. We will use the Multivariate Delta Method. We know that $\sqrt{n}\left(\frac{\overline{X}_1}{\overline{X}_2} - \frac{\mu_1}{\mu_2}\right) \xrightarrow{\text{d}} N(0, \nabla_\mu^T \Sigma \nabla_\mu)$.

By the Central Limit Theorem, we know that

$$\sqrt{n}\begin{pmatrix} \overline{X}_1 - \mu_1 \\ \overline{X}_2 - \mu_2 \end{pmatrix} \xrightarrow{\text{d}} N(0, \Sigma).$$

Note that $Y_n = g(\overline{X}_1, \overline{X_2})$ where $g(s_1, s_2) = \frac{s_1}{s_2}$. Thus

$$\nabla g(s) = \begin{pmatrix} \frac{\partial g}{\partial s_1} \\ \frac{\partial g}{\partial s_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{s_2} \\ -\frac{s_1}{s_2^2} \end{pmatrix}$$

and so

$$\begin{aligned} \nabla_\mu^T \Sigma \nabla_\mu &= \begin{pmatrix} \frac{1}{\mu_2} \\ -\frac{\mu_1}{\mu_2^2} \end{pmatrix}^T \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ -\frac{\mu_1}{\mu_2^2} \end{pmatrix} \\ &= \frac{\mu_2^2 \sigma_{11} - 2\mu_1\mu_2\sigma_{12} + \mu_1^2\sigma_{22}}{\mu_2^4}. \end{aligned}$$

It follows that Y_n has limiting distribution $N\left(\frac{\mu_1}{\mu_2}, \frac{\mu_2^2\sigma_{11} - 2\mu_1\mu_2\sigma_{12} + \mu_1^2\sigma_{22}}{n\mu_2^4}\right)$. $\square$

Problem 5.16. Construct an example where $X_n \xrightarrow{\text{d}} X$ and $Y_n \xrightarrow{\text{d}} Y$ but $X_n + Y_n$ does not converge in distribution to $X + Y$.

Solution. Let $X \sim N(0,1)$, and let $X_n = (-1)^n \cdot X$ and $Y_n = (-1)^{n+1}$ for all positive integers n.

All the X_n and Y_n have the same distribution function as X, so $X_n \xrightarrow{\text{d}} X$ and $Y_n \xrightarrow{\text{d}} X$. But $X_n + Y_n = (-1)^n X + (-1)^{n+1} X = 0$ for all n, so $X_n + Y_n \xrightarrow{\text{d}} 0$. Certainly $2X$ and 0 are not the same, and we are done. $\square$