Probability and Statistics MA6.101

Homework 8

Topics Covered: Convergence, Central Limit Theorem

Q1: Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent random variables such that

$$X_n = \begin{cases} 0 & \text{with probability } 1 - \dfrac{1}{n} \\ 1 & \text{with probability } \dfrac{1}{n} \end{cases}$$

(a) Does this sequence converge to 0 in mean square? That is, does $X_n \xrightarrow{m.s.} 0$?

(b) Does this sequence converge to 0 almost surely? That is, does $X_n \xrightarrow{a.s.} 0$?

Solution:

(a) **Mean square convergence:**
For $X_n \xrightarrow{m.s.} 0$, we need $\mathbb{E}[X_n^2] \to 0$ as $n \to \infty$.
Since $X_n \in \{0, 1\}$, we have $X_n^2 = X_n$. Therefore:

$$\begin{aligned} \mathbb{E}[X_n^2] &= \mathbb{E}[X_n] \\ &= 0 \cdot \left(1 - \frac{1}{n}\right) + 1 \cdot \frac{1}{n} \\ &= \frac{1}{n} \to 0 \text{ as } n \to \infty \end{aligned}$$

Therefore, $X_n \xrightarrow{m.s.} 0$.

(b) **Almost sure convergence:**
For $X_n \xrightarrow{a.s.} 0$, we need to check if $\mathbb{P}\left(\lim_{n\to\infty} X_n = 0\right) = 1$.
An equivalent condition is: for any $\epsilon > 0$,

$$\mathbb{P}(|X_n - 0| < \epsilon \text{ for all } n \geq m) \to 1 \text{ as } m \to \infty$$

For $0 < \epsilon < 1$, we have $|X_n - 0| < \epsilon$ if and only if $X_n = 0$.
Therefore, for any m:

$$\begin{aligned} \mathbb{P}\{|X_n - 0| < \epsilon \text{ for all } n \geq m\} &= \lim_{n\to\infty} \prod_{i=m}^{n} \mathbb{P}(X_i = 0) \\ &= \lim_{n\to\infty} \prod_{i=m}^{n} \left(1 - \frac{1}{i}\right) \\ &= \lim_{n\to\infty} \prod_{i=m}^{n} \frac{i-1}{i} \\ &= \lim_{n\to\infty} \frac{(m-1)}{m} \cdot \frac{m}{(m+1)} \cdots \frac{(n-1)}{n} \\ &= \lim_{n\to\infty} \frac{m-1}{n} \to 0 \neq 1 \end{aligned}$$

Since this probability goes to 0 (not 1) as $n \to \infty$, the sequence does not converge to 0 almost surely.
Therefore, $X_n \overset{a.s.}{\not\longrightarrow} 0$.

Q2: A company measures how long each customer service call lasts. Historically, individual call durations have mean $\mu = 10$ minutes and standard deviation $\sigma = 10$ minutes, and call durations are roughly exponential.

(a) For a random sample of 50 calls, what is the probability that the sample average call duration exceeds 12 minutes?

(b) What is the minimum sample size n needed so that the probability the sample average exceeds 12 minutes is at most 0.05?

(c) For a sample of 80 calls, what is the probability the sample average lies between 9 minutes and 11 minutes?

Use:

$$\Phi^{-1}(0.95) = 1.6449, \qquad \Phi(\sqrt{2}) = 0.9214, \qquad \Phi(0.8944) = 0.8147.$$

Solution:

Define $T_1, T_2, \ldots$ as independent call durations with mean $\mu = 10$ and standard deviation $\sigma = 10$. For sample size n define $\overline{T}_n = \frac{1}{n}\sum_{i=1}^{n} T_i$. By the Central Limit Theorem,

$$\overline{T}_n \approx N\Big(\mu, \frac{\sigma}{\sqrt{n}}\Big) = N\Big(10, \frac{10}{\sqrt{n}}\Big),$$

so for any threshold a,

$$P(\overline{T}_n > a) \approx P\Big(Z > \frac{a-\mu}{\sigma/\sqrt{n}}\Big), \quad Z \sim N(0,1).$$

(a) $n = 50$, find $P(\overline{T}_{50} > 12)$.

$$\begin{aligned} P(\overline{T}_{50} > 12) &\approx P\Big(Z > \frac{12-\mu}{\sigma/\sqrt{50}}\Big) = P\Big(Z > \frac{12-10}{10/\sqrt{50}}\Big) \\ &= P\Big(Z > \frac{2}{10/\sqrt{50}}\Big) = P\Big(Z > \frac{2\sqrt{50}}{10}\Big) = P\Big(Z > \frac{\sqrt{50}}{5}\Big). = P(Z > \sqrt{2}) \end{aligned}$$

$$P(\overline{T}_{50} > 12) \approx 1 - \Phi(\sqrt{2}) = 1 - 0.9214 = \boxed{0.0786}.$$

(b) Minimum integer n so $P(\overline{T}_n > 12) \le 0.05$.

We require

$$1 - \Phi\Big(\frac{12-\mu}{\sigma/\sqrt{n}}\Big) \le 0.05 \quad \Longleftrightarrow \quad \Phi\Big(\frac{12-\mu}{\sigma/\sqrt{n}}\Big) \ge 0.95.$$

Let $z_{0.95} = \Phi^{-1}(0.95) = 1.6449$. Then

$$\frac{12-10}{10/\sqrt{n}} \ge z_{0.95} \quad \Longrightarrow \quad \frac{2\sqrt{n}}{10} \ge 1.6449 \quad \Longrightarrow \quad \frac{\sqrt{n}}{5} \ge 1.6449.$$

So

$$\sqrt{n} \ge 5 \times 1.6449 = 8.2245 \quad \Longrightarrow \quad n \ge (8.2245)^2 \approx 67.6425.$$

Hence the smallest integer n is $\boxed{68}$.

(c) $n = 80$**, find** $P(9 \le \overline{T}_{80} \le 11)$**.**

$$\begin{aligned} P(9 \le \overline{T}_{80} \le 11) &\approx P\Big(\frac{9-\mu}{\sigma/\sqrt{80}} \le Z \le \frac{11-\mu}{\sigma/\sqrt{80}}\Big) \\ &= P\Big(-\frac{\sqrt{80}}{10} \le Z \le \frac{\sqrt{80}}{10}\Big) = \Phi\Big(\frac{\sqrt{80}}{10}\Big) - \Phi\Big(-\frac{\sqrt{80}}{10}\Big) \\ &= 2\Phi\Big(\frac{\sqrt{80}}{10}\Big) - 1. \end{aligned}$$

Note $\dfrac{\sqrt{80}}{10} \approx 0.8944$ and $\Phi(0.8944) = 0.8147$. Thus

$$P(9 \le \overline{T}_{80} \le 11) \approx 2(0.8147) - 1 = \boxed{0.6294}.$$

Q3: A factory produces metal rods whose individual lengths historically have mean $\mu = 100$ mm and standard deviation $\sigma = 5$ mm, and lengths are approximately normal.

(a) For a single randomly chosen rod, what is the probability its length exceeds 110 mm?

(b) If a sample of 36 rods is taken, what is the probability that the sample mean length exceeds 103 mm?

Use:

$$\Phi(2) = 0.9772, \quad \Phi(3.6) = 0.9998$$

Solution:

Define X as the length of one rod. Then $X \sim N(100, 5^2)$. For a sample of size n, $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\Big(100, \frac{5}{\sqrt{n}}\Big)$.

(a) Single rod: $P(X > 110)$

$$P(X > 110) = P\Big(\frac{X-\mu}{\sigma} > \frac{110-100}{5}\Big) = P(Z > 2).$$

Using $\Phi(2) = 0.9772$,

$$P(X > 110) = 1 - 0.9772 = \boxed{0.0228}.$$

(b) Sample of $n = 36$: $P(\overline{X}_{36} > 103)$

$$P(\overline{X}_{36} > 103) = P\Big(Z > \frac{103 - 100}{5/\sqrt{36}}\Big) = P\Big(Z > \frac{3\sqrt{36}}{5}\Big) = P(Z > 3.6).$$

With $\Phi(3.6) = 0.9998$,

$$P(\overline{X}_{36} > 103) = 1 - 0.9998 = \boxed{0.0002}.$$

Q4: Let $X_n \sim \text{Uniform}\left(0, \frac{1}{n}\right)$ for $n = 1, 2, 3, \ldots$

Show that $X_n \xrightarrow{p} 0$ (i.e., X_n converges in probability to 0).

Solution:

Recall the definition of convergence in probability: $X_n \xrightarrow{p} X$ if

$$\lim_{n\to\infty} \mathbb{P}\big(|X_n - X| \geq \varepsilon\big) = 0 \quad \text{for all } \varepsilon > 0$$

Here $X = 0$, so we need to show that for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - 0| \geq \varepsilon) = 0.$$

Since $X_n \geq 0$ we have $|X_n - 0| = X_n$. Therefore, we need to compute $\mathbb{P}(X_n \geq \varepsilon)$.

Since $X_n \sim \text{Uniform}\left(0, \frac{1}{n}\right)$, the probability density function is

$$f_{X_n}(x) = \begin{cases} n, & \text{if } x \in \left[0, \frac{1}{n}\right], \\ 0, & \text{otherwise.} \end{cases}$$

We must consider two cases depending on the relationship between ε and $\frac{1}{n}$.

Case A: $\frac{1}{n} < \varepsilon$

In this case, the entire support of X_n is $\left[0, \frac{1}{n}\right] \subset [0, \varepsilon)$. This means every possible value of X_n is strictly less than ε. Therefore, the event $\{X_n \geq \varepsilon\}$ is impossible, so

$$\mathbb{P}(X_n \geq \varepsilon) = 0.$$

Case B: $\frac{1}{n} \geq \varepsilon$

In this case, the interval $[\varepsilon, 1/n]$ is non-empty and lies within the support of X_n. We compute

$$\begin{aligned} \mathbb{P}(X_n \geq \varepsilon) &= \int_{\varepsilon}^{1/n} n\,dx \\ &= n\left[\frac{1}{n} - \varepsilon\right] \\ &= 1 - n\varepsilon. \end{aligned}$$

Thus, for any fixed $\varepsilon > 0$ and any $n \in \mathbb{N}$,

$$\mathbb{P}(X_n \geq \varepsilon) = \begin{cases} 0, & \text{if } \frac{1}{n} < \varepsilon, \\ 1 - n\varepsilon, & \text{if } \frac{1}{n} \geq \varepsilon. \end{cases}$$

Now fix an arbitrary $\varepsilon > 0$. Choose $N = \lfloor \frac{1}{\varepsilon} \rfloor + 1$. Then for all $n \geq N$,

$$\frac{1}{n} \leq \frac{1}{N} < \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

This means that for all $n \geq N$, we are in Case A, so $\mathbb{P}(X_n \geq \varepsilon) = 0$. Therefore,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - 0| \geq \varepsilon) = 0.$$

Since this holds for every $\varepsilon > 0$, we have proven that $X_n \xrightarrow{p} 0$.

Q5: Let X be any random variable. For each positive integer n, let $Y_n \sim N(0, 1/n)$ independent of X and set $X_n = X + Y_n$. Show that X_n converges to X in mean-square:

$$X_n \xrightarrow{m.s.} X.$$

Solution :

We need to show that $\mathbb{E}\big(|X_n - X|^2\big) \to 0$ as $n \to \infty$.

$$\mathbb{E}\big(|X_n - X|^2\big) = \mathbb{E}\big(|(X + Y_n) - X|^2\big) = \mathbb{E}\big(|Y_n|^2\big).$$

Thus, we only need to evaluate $\mathbb{E}(Y_n^2)$.

Because $Y_n \sim N(0, 1/n)$, its variance is $\mathrm{Var}(Y_n) = 1/n$ and mean $\mathbb{E}[Y_n] = 0$. For any random variable,

$$\mathbb{E}(Y_n^2) = \mathrm{Var}(Y_n) + (\mathbb{E}[Y_n])^2.$$

Hence,

$$\mathbb{E}(Y_n^2) = \frac{1}{n} + 0^2 = \frac{1}{n}.$$

$$\mathbb{E}\big(|X_n - X|^2\big) = \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

Therefore, by the definition of mean-square convergence,

$$\boxed{X_n \xrightarrow{m.s.} X.}$$

Q6: For each $n \geq 1$, let the random variable X_n be defined as

$$P(X_n = 0) = 1 - \frac{1}{n}, \qquad P(X_n = n) = \frac{1}{n}.$$

Show that X_n converges in distribution to the degenerate random variable X that equals 0 with probability 1. That is,

$$X_n \xrightarrow{d} X, \qquad X \equiv 0.$$

Solution:

$$X_n \xrightarrow{d} X \quad \Longleftrightarrow \quad \lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{for all } x \text{ where } F_X(x) \text{ is continuous.}$$

Since X is degenerate at 0, its CDF is:

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

Note that $F_X(x)$ is continuous at all $x \neq 0$.

The CDF of X_n:

Because X_n takes only the two values 0 and n:

$$F_{X_n}(x) = P(X_n \leq x) = \begin{cases} 0, & x < 0, \\ 1 - \dfrac{1}{n}, & 0 \leq x < n, \\ 1, & x \geq n. \end{cases}$$

Take pointwise limits at continuity points of F_X.

Case A: $x < 0$ For all n, $F_{X_n}(x) = 0$. Hence,

$$\lim_{n\to\infty} F_{X_n}(x) = 0 = F_X(x).$$

Case B: $x > 0$ For any fixed $x > 0$, when n is large enough ($n > x$), we have $0 \leq x < n$, so

$$F_{X_n}(x) = 1 - \frac{1}{n}.$$

Thus,

$$\lim_{n\to\infty} F_{X_n}(x) = 1 = F_X(x).$$

We do not check $x = 0$ since F_X is discontinuous there.

Since $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ for every x where F_X is continuous, we conclude

$$\boxed{X_n \xrightarrow{d} X, \quad X \equiv 0.}$$

Q7: Let $X_n = \frac{1}{\sqrt{n}} Z$, where $\mathbb{E}(Z^2) < \infty$. Show that $X_n \to 0$ in distribution.

Solution: Proving convergence in mean square is easier. We have

$$\mathbb{E}(X_n^2) = \mathbb{E}\left(\left(\frac{1}{\sqrt{n}} Z\right)^2\right) = \frac{1}{n}\mathbb{E}(Z^2) \to 0 \text{ as } n \to \infty.$$

Thus, $X_n \xrightarrow{m.s.} 0$ and hence $X_n \xrightarrow{d} 0$. Because convergence in mean square implies convergence in distribution.

Q8: If $X_n \xrightarrow{L^r} X$ for some $r \geq 1$, then show that $X_n \xrightarrow{p} X$.

Solution: For any $\epsilon > 0$, we have

$$P(|X_n - X| \geq \epsilon) = P(|X_n - X|^r \geq \epsilon^r) \leq \frac{E|X_n - X|^r}{\epsilon^r} \quad \text{(by Markov's inequality, since } r \geq 1\text{)}.$$

Since by assumption $\lim_{n\to\infty} E(|X_n - X|^r) = 0$, we conclude

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0, \quad \text{for all } \epsilon > 0.$$

Q9: Before starting to play roulette in a casino you want to look for biases you can exploit. You therefore watch 100 independent rounds; each round yields a number in $\{1, \ldots, 36\}$, and you count the number of rounds for which the outcome is *odd*. If the count exceeds 55 you decide that the roulette is not fair.

Assuming the roulette is fair (so that odd and even are equally likely), find an approximation for the probability that you will make the wrong decision (i.e. the probability of deciding "not fair" when in fact the wheel is fair).

Remark / clarification about roulette: A standard (idealized) roulette without the green zero pocket would produce numbers 1 to 36 with equal probability; half of these are odd and half even. In that idealized model a *fair* wheel gives $\Pr(\text{odd}) = \frac{1}{2}$. The test described checks whether the observed count of odd outcomes in 100 spins is unusually large (strictly greater than 55).

Solution Let X denote the number of odd outcomes in the 100 independent spins. Under the null hypothesis that the wheel is fair,

$$X \sim \text{Binomial}\left(n = 100,\ p = \tfrac{1}{2}\right).$$

Hence

$$\mathbb{E}[X] = np = 100 \cdot \tfrac{1}{2} = 50, \qquad \text{Var}(X) = np(1-p) = 100 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = 25,$$

so the standard deviation is $\sigma_X = \sqrt{25} = 5$.

We must approximate the Type I error probability

$$\Pr(\text{wrong decision}) = \Pr(X > 55) = \Pr(X \geq 56).$$

Normal (CLT) approximation without continuity correction. By the Central Limit Theorem we approximate X by a normal random variable with mean 50 and variance 25. Thus

$$\Pr(X > 55) \approx \Pr\Big(Z > \frac{55-50}{5}\Big) = \Pr(Z > 1),$$

where $Z \sim N(0,1)$. Using standard normal tables,

$$\Pr(Z > 1) = 1 - \Phi(1) \approx 1 - 0.84134 = 0.15866.$$

Normal approximation with continuity correction (more accurate). Using the continuity correction, $\Pr(X \geq 56) = \Pr(X > 55.5)$ is approximated by

$$\Pr\Big(Z > \frac{55.5-50}{5}\Big) = \Pr(Z > 1.1) = 1 - \Phi(1.1) \approx 1 - 0.86433 = 0.13567.$$

Answer (approximate): The probability of wrongly declaring the fair roulette "not fair" (Type I error) is about

$$\boxed{\Pr(X > 55) \approx 0.14 \text{ to } 0.16.}$$

Remark (exact expression). The exact probability under the binomial model is

$$\Pr(X \geq 56) = \sum_{k=56}^{100} \binom{100}{k} \left(\tfrac{1}{2}\right)^{100},$$

Q10: We have a bag with n blue balls and n red balls, where $n \geq 10$. We randomly draw 10 balls without replacement. Let X_n be the number of blue balls drawn. Prove that:

$$X_n \xrightarrow{d} \text{Binomial}(10, 1/2)$$

A: The experiment consists of drawing a sample of size $k = 10$ from a finite population of size $N = 2n$ without replacement. The population contains $K = n$ successes (blue balls).
The probability mass function (PMF) of X_n for $j \in \{0, 1, \ldots, 10\}$ is:

$$p_{X_n}(j) = \frac{\binom{n}{j}\binom{n}{10-j}}{\binom{2n}{10}}$$

To prove convergence in distribution ($X_n \xrightarrow{d} Y$), we show that $\lim_{n\to\infty} P(X_n = j) = P(Y = j)$. We evaluate the limit:

$$\lim_{n\to\infty} \frac{\binom{n}{j}\binom{n}{10-j}}{\binom{2n}{10}}$$

First, we expand the binomial coefficients:

$$P(X_n = j) = \frac{\frac{n(n-1)\cdots(n-j+1)}{j!} \cdot \frac{n(n-1)\cdots(n-10+j+1)}{(10-j)!}}{\frac{2n(2n-1)\cdots(2n-9)}{10!}}$$

Rearranging the expression:

$$P(X_n = j) = \frac{10!}{j!(10-j)!} \cdot \frac{[n(n-1)\cdots(n-j+1)] \cdot [n(n-1)\cdots(n-10+j+1)]}{2n(2n-1)\cdots(2n-9)}$$

$$P(X_n = j) = \binom{10}{j} \cdot \frac{\left(\prod_{i=0}^{j-1}(n-i)\right)\left(\prod_{i=0}^{9-j}(n-i)\right)}{\prod_{i=0}^{9}(2n-i)}$$

The fraction is a ratio of two polynomials in n. The numerator is a polynomial of degree $j + (10 - j) = 10$, with leading term n^{10}. The denominator is a polynomial of degree 10, with leading term $(2n)^{10} = 2^{10}n^{10}$.

The limit of a rational function where the degrees of the numerator and denominator are equal is the ratio of their leading coefficients.

$$\lim_{n\to\infty} \frac{\left(\prod_{i=0}^{j-1}(n-i)\right)\left(\prod_{i=0}^{9-j}(n-i)\right)}{\prod_{i=0}^{9}(2n-i)} = \frac{1}{2^{10}}$$

Therefore,

$$\lim_{n\to\infty} P(X_n = j) = \binom{10}{j} \cdot \frac{1}{2^{10}}$$

This is the PMF of a Binomial$(10, 1/2)$ distribution.

Q11: Let $Y_1, Y_2, \ldots$ be independent random variables, where $Y_n \sim$ Bernoulli $\left(\frac{n}{n+1}\right)$ for $n = 1, 2, 3, \ldots$. We define the sequence $\{X_n, n = 2, 3, 4, \ldots\}$ as

$$X_{n+1} = Y_1 Y_2 Y_3 \cdots Y_n, \quad \text{for } n = 1, 2, 3, \ldots$$

Show that $X_n \xrightarrow{\text{a.s.}} 0$.

A: To show that X_n converges almost surely to 0, we must prove that the probability of the event $\{\lim_{n\to\infty} X_n = 0\}$ is 1.

$$P\left(\lim_{n\to\infty} X_n = 0\right) = 1$$

The random variable $X_n = \prod_{k=1}^{n-1} Y_k$ is a product of Bernoulli variables. As such, X_n can only take values in $\{0, 1\}$.

The sequence X_n converges to 0 if at least one of the variables Y_k is 0. The only way for the limit not to be 0 is if $X_n = 1$ for all n. This occurs if and only if $Y_k = 1$ for all $k \geq 1$.

Let's consider the complementary event, where the limit is not 0.

$$\left\{\lim_{n\to\infty} X_n \neq 0\right\} \iff \{X_n = 1 \text{ for all } n \geq 2\} \iff \{Y_k = 1 \text{ for all } k \geq 1\}$$

We can now calculate the probability of this event. Since the Y_k are independent:

$$
\begin{aligned}
P\left(\lim_{n\to\infty} X_n \neq 0\right) &= P(Y_k = 1 \text{ for all } k \geq 1) \\
&= P\left(\bigcap_{k=1}^{\infty}\{Y_k = 1\}\right) \\
&= \lim_{t\to\infty} P\left(\bigcap_{k=1}^{t}\{Y_k = 1\}\right) \\
&= \lim_{t\to\infty} \prod_{k=1}^{t} P(Y_k = 1)
\end{aligned}
$$

We are given that $P(Y_k = 1) = \frac{k}{k+1}$. Substituting this into the product gives:

$$\lim_{t\to\infty} \prod_{k=1}^{t} \frac{k}{k+1}$$

This is a telescoping product:

$$\prod_{k=1}^{t} \frac{k}{k+1} = \left(\frac{1}{2}\right)\cdot\left(\frac{2}{3}\right)\cdot\left(\frac{3}{4}\right)\cdots\left(\frac{t}{t+1}\right) = \frac{1}{t+1}$$

Taking the limit as $t \to \infty$:

$$\lim_{t\to\infty} \frac{1}{t+1} = 0$$

Thus, we have shown that the probability of the complementary event is 0:

$$P\left(\lim_{n\to\infty} X_n \neq 0\right) = 0$$

Therefore, the probability of our original event must be 1:

$$P\left(\lim_{n\to\infty} X_n = 0\right) = 1 - 0 = 1$$

This satisfies the definition of almost sure convergence.

Q12: Let $Y_1, Y_2, Y_3, \dots$ be a sequence of i.i.d. random variables with mean $E[Y_i] = \mu$ and finite variance $Var(Y_i) = \sigma^2$. Define the sequence $\{X_n, n = 2, 3, \dots\}$ as

$$X_n = \frac{Y_1Y_2 + Y_2Y_3 + \cdots + Y_{n-1}Y_n + Y_nY_1}{n}, \quad \text{for } n = 2, 3, \dots$$

Show that $X_n \xrightarrow{p} \mu^2$.

A: We will prove convergence in probability by using Chebyshev's Inequality. The strategy is to show that $E[X_n] = \mu^2$ and $\lim_{n\to\infty} Var(X_n) = 0$.

Let us define a sequence of random variables $Z_i = Y_iY_{i+1}$ for $i = 1, \dots, n-1$ and $Z_n = Y_nY_1$. Then $X_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$.

Since the Y_i are independent and identically distributed (i.i.d.), the expectation of each Z_i is:

$$E[Z_i] = E[Y_i Y_{i+1}] = E[Y_i]E[Y_{i+1}] = \mu \cdot \mu = \mu^2$$

This holds for all i, including $Z_n = Y_n Y_1$.
By the linearity of expectation, the expected value of X_n is:

$$E[X_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} Z_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[Z_i] = \frac{1}{n}\sum_{i=1}^{n} \mu^2 = \frac{n\mu^2}{n} = \mu^2$$

The variance of X_n is given by:

$$Var(X_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} Z_i\right)$$

The variance of the sum is the sum of the variances and covariances:

$$Var\left(\sum_{i=1}^{n} Z_i\right) = \sum_{i=1}^{n} Var(Z_i) + \sum_{i \neq j} Cov(Z_i, Z_j)$$

We need to calculate $Var(Z_i)$ and $Cov(Z_i, Z_j)$. First, we note that $E[Y_i^2] = Var(Y_i) + (E[Y_i])^2 = \sigma^2 + \mu^2$.
Variance of Z_i:

$$\begin{aligned} Var(Z_i) &= E[Z_i^2] - (E[Z_i])^2 = E[(Y_i Y_{i+1})^2] - (\mu^2)^2 \\ &= E[Y_i^2 Y_{i+1}^2] - \mu^4 \\ &= E[Y_i^2]E[Y_{i+1}^2] - \mu^4 \quad \text{(by independence)} \\ &= (\sigma^2 + \mu^2)(\sigma^2 + \mu^2) - \mu^4 = (\sigma^2 + \mu^2)^2 - \mu^4 \end{aligned}$$

This is a finite constant.
Covariance of Z_i, Z_j: The terms $Z_i = Y_i Y_{i+1}$ are not independent if they share a common variable.

- **Case 1: Non-adjacent terms.** If the indices are such that $\{i, i+1\} \cap \{j, j+1\} = \emptyset$, then the variables $Y_i, Y_{i+1}, Y_j, Y_{j+1}$ are all independent.

  $$E[Z_i Z_j] = E[Y_i Y_{i+1} Y_j Y_{j+1}] = E[Y_i]E[Y_{i+1}]E[Y_j]E[Y_{j+1}] = \mu^4$$

  So, $Cov(Z_i, Z_j) = E[Z_i Z_j] - E[Z_i]E[Z_j] = \mu^4 - \mu^2\mu^2 = 0$.
- **Case 2: Adjacent terms.** Consider $Cov(Z_i, Z_{i+1})$ (indices are modulo n).

  $$\begin{aligned} E[Z_i Z_{i+1}] &= E[Y_i Y_{i+1} Y_{i+1} Y_{i+2}] = E[Y_i Y_{i+1}^2 Y_{i+2}] \\ &= E[Y_i]E[Y_{i+1}^2]E[Y_{i+2}] \quad \text{(by independence)} \\ &= \mu(\sigma^2 + \mu^2)\mu = \mu^2(\sigma^2 + \mu^2) \end{aligned}$$

  So, $Cov(Z_i, Z_{i+1}) = \mu^2(\sigma^2 + \mu^2) - \mu^4 = \mu^2\sigma^2$. This is a finite constant.

The only non-zero covariances are between adjacent terms. For each Z_i, there are two neighbors (Z_{i-1} and Z_{i+1}), leading to $2n$ non-zero covariance terms.

$$\begin{aligned}
Var\left(\sum_{i=1}^{n} Z_i\right) &= n \cdot Var(Z_1) + 2n \cdot Cov(Z_1, Z_2) \\
&= n\left(((\sigma^2+\mu^2)^2 - \mu^4) + 2\mu^2\sigma^2\right)
\end{aligned}$$

The term in the parenthesis is a finite constant. Thus, the variance of the sum is of order $O(n)$. Plugging this back into the expression for $Var(X_n)$:

$$Var(X_n) = \frac{1}{n^2}\left[n\left(Var(Z_1) + 2Cov(Z_1, Z_2)\right)\right] = \frac{Var(Z_1) + 2Cov(Z_1, Z_2)}{n}$$

As $n \to \infty$, the variance approaches zero:

$$\lim_{n\to\infty} Var(X_n) = \lim_{n\to\infty} \frac{\text{Constant}}{n} = 0$$

Chebyshev's inequality states that for any $\epsilon > 0$:

$$P(|X_n - E[X_n]| \geq \epsilon) \leq \frac{Var(X_n)}{\epsilon^2}$$

Substituting our results:

$$P(|X_n - \mu^2| \geq \epsilon) \leq \frac{Var(Z_1) + 2Cov(Z_1, Z_2)}{n\epsilon^2}$$

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

$$\lim_{n\to\infty} P(|X_n - \mu^2| \geq \epsilon) \leq \lim_{n\to\infty} \frac{\text{Constant}}{n\epsilon^2} = 0$$

Since probability cannot be negative, this limit must be exactly 0. This is the definition of convergence in probability.
We have shown that $\lim_{n\to\infty} P(|X_n - \mu^2| \geq \epsilon) = 0$. Therefore,

$$X_n \xrightarrow{p} \mu^2$$

Q13: [**Bonus**] Show that in the limit of large N, the binomial distribution of n out of N objects becomes a Gaussian distribution. [Hint: Write $P(n) = \binom{N}{n} p^n q^{N-n}$. Take logarithms of both sides and use Stirling's approximation $\ln(N!) \approx N\ln(N) - N$. Then apply Taylor's series around the mean of the distribution].

A:
We begin with the binomial distribution for the probability of n successes in N trials, where p is the success probability and $q = 1 - p$:

$$P(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

To manage the factorials, we take the natural logarithm and apply Stirling's approximation, $\ln(k!) \approx k \ln k - k$.

$$\ln P(n) = \ln(N!) - \ln(n!) - \ln((N-n)!) + n \ln p + (N-n) \ln q$$

$$\ln P(n) \approx (N \ln N - N) - (n \ln n - n) - ((N-n) \ln(N-n) - (N-n)) + n \ln p + (N-n) \ln q$$

The linear terms in N and n cancel out, leaving:

$$\ln P(n) \approx N \ln N - n \ln n - (N-n) \ln(N-n) + n \ln p + (N-n) \ln q$$

We now analyze the distribution near its peak, the mean $\mu = Np$, by considering a small deviation x, such that $n = Np + x$ and thus $N - n = Nq - x$. We expand the logarithmic terms using the Taylor series $\ln(1+y) \approx y - y^2/2$. For example:

$$\ln(Np + x) = \ln(Np) + \ln\left(1 + \frac{x}{Np}\right) \approx \ln(Np) + \frac{x}{Np} - \frac{x^2}{2(Np)^2}$$

Similarly for $\ln(Nq - x)$:

$$\ln(Nq - x) \approx \ln(Nq) - \frac{x}{Nq} - \frac{x^2}{2(Nq)^2}$$

Substituting these expansions into the expression for $\ln P(n)$ and collecting terms by powers of x, we find that all constant terms and all terms linear in x cancel out. The only significant remaining terms are quadratic in x:

$$\begin{aligned}
\ln P(n) &\approx -\frac{x^2}{2Np} - \frac{x^2}{2Nq} \\
&= -\frac{x^2}{2N}\left(\frac{1}{p} + \frac{1}{q}\right) \\
&= -\frac{x^2}{2N}\left(\frac{q+p}{pq}\right) = -\frac{x^2}{2Npq}
\end{aligned}$$

Exponentiating both sides and substituting back $x = n - Np$, we get the final form:

$$P(n) \approx A \cdot \exp\left(-\frac{(n-Np)^2}{2Npq}\right)$$

This is the equation for a Gaussian distribution with a mean $\mu = Np$ and a variance $\sigma^2 = Npq$, where A is a normalization constant. Thus, the binomial distribution becomes Gaussian for large N.