Convergent Random Variables

Christopher Hundt / Supervised by Prof. David Wolfson / christopher.hundt@mail.mcgill.ca

Introduction

What is convergence of a sequence of random variables?

If X_n is a random variable for every $n = 1, 2, \ldots$ and Y is another random variable then we say that X_n converges to Y in probability, written $X_n \to Y$ in probability, if for every $\varepsilon > 0$ and $\delta > 0$ there exists some N such that $\Pr(|Y - X_n| > \delta) < \varepsilon$ for all $n > N$. The interpretation of this is that the X_n are ‘‘behaving’’ more and more like Y.

Consider this example: let X_n be such that $P(X_n = 1) = 1/n$ and $P(X_n = 0) = 1 - 1/n$. Then $X_n \to Y$ in probability, where $Y = 0$ with probability 1, because, for any $\varepsilon > 0$, $\Pr(|X_n - Y| > \varepsilon) = \Pr(X_n \neq 0) = 1/n$ becomes arbitrarily small as n increases.

What are independent random variables?

Two random variables X and Y are independent if, for all sets A and B, $\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B)$. The intuition here is that X and Y have nothing to do with each other: knowing the value of X at a particular time tells you nothing about what the value of Y is at that same time. If two random variables are not independent, they are called dependent.

Do the apparently unrelated notions of convergence and independence have any connection?

Upon learning of convergent random variables, I wondered whether convergence implies some sort of dependence: if $X_n \to Y$ in probability, then must each X_n be dependent on Y eventually? The definition of convergence makes no reference to independence or dependence. On the other hand, the intuitive ideas that X_n converging to Y means the X_n eventually ‘‘look like’’ Y and that random variables are independent only if their values have ‘‘nothing to do with each other’’ suggests a relationship.

The literature on random variables did not seem to consider this question. So Prof. Wolfson suggested that I investigate. It turned out that the intuition is indeed correct!

Main result

THEOREM: Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables. Let Y be a non-degenerate random variable. If $X_n \to Y$ in probability then there exists some N such that X_n is not independent of Y for all $n > N$.

I then asked myself: is there a relationship between correlation and convergence?

Correlation

What is correlation?

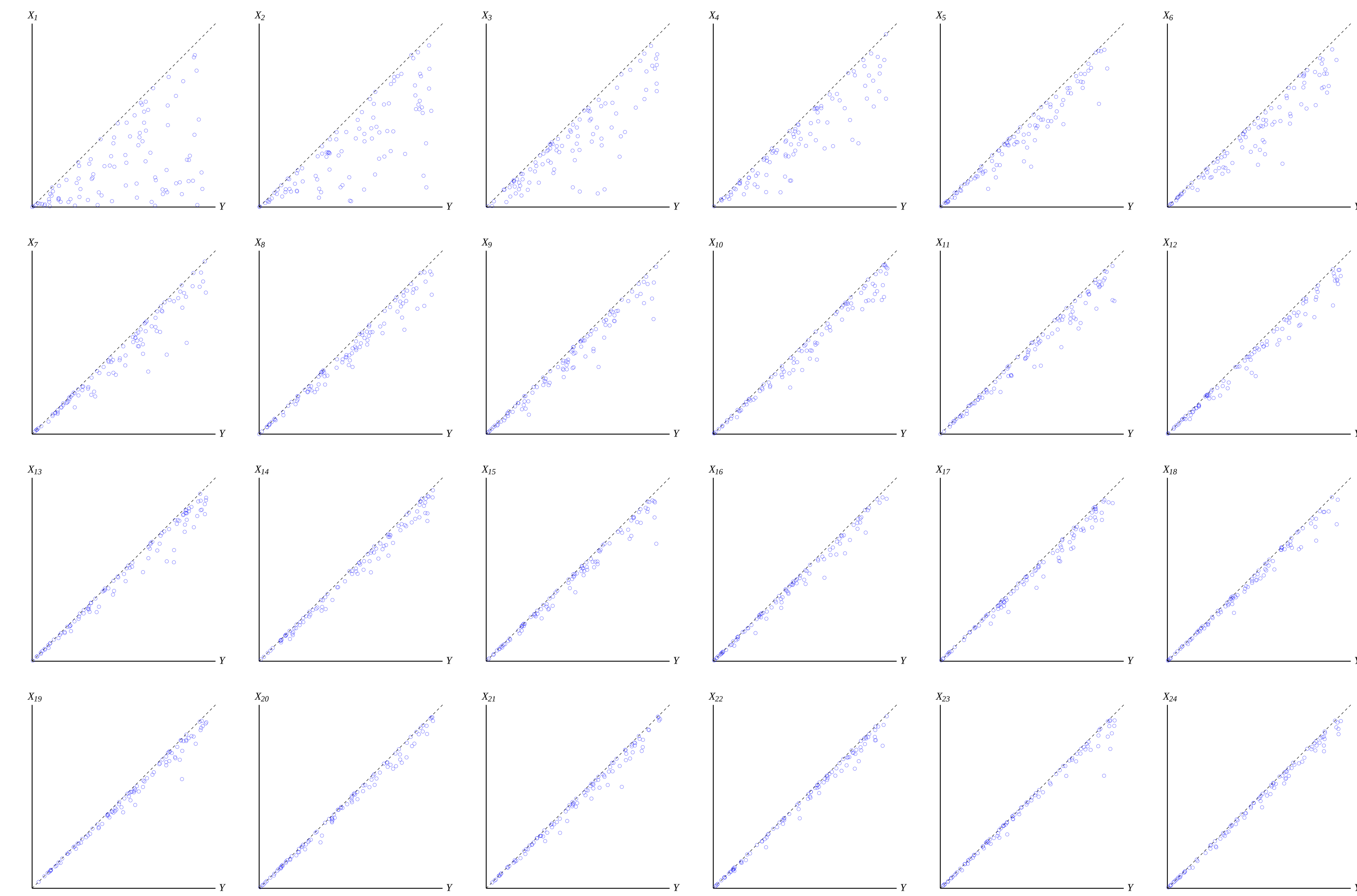
If X and Y are two random variables then we define the covariance between X and Y as $\mathrm{Cov}(X,Y) = E[(X - \sigma_X)(Y - \sigma_Y)]$, where σ_X and σ_Y are the standard deviations of X and Y, respectively. Then we define the correlation between X and Y as $\rho(X,Y) = \mathrm{Cov}(X,Y)/(\sigma_X \sigma_Y)$. The correlation is intuitively thought of as a measure of linear dependence between X and Y: if $|\rho(X,Y)| = 1$ then X and Y are completely linearly dependent, and if $\rho(X,Y) = 0$ then there is no linear relationship between X and Y.

Is correlation related to convergence?

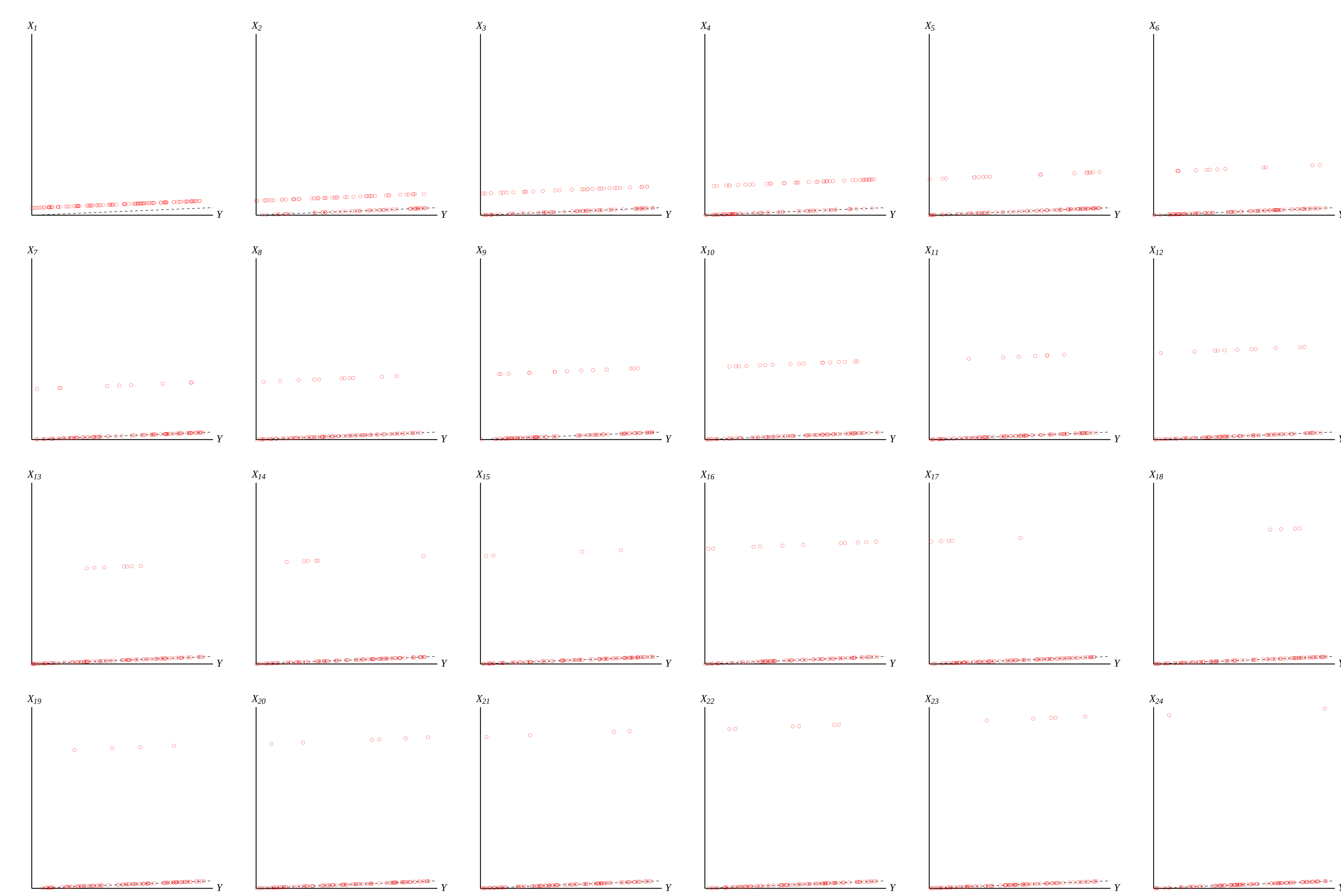
Now we wondered what the nature of the dependence imposed by convergence is. Does a sequence of random variables become more and more correlated with the limiting random variable? The answer is (usually) yes!

THEOREM: Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of bounded random variables. Let Y be a non-degenerate random variable. If $X_n \to Y$ in probability then $\rho(X_n, Y) \to 1$ as $n \to \infty$.

Consider an example of this theorem in action: Let Y be a uniform(0,1) random variable and X_n defined as $X_n = Y \cdot Z_n$, where $Z_n = \max(U_1, \ldots, U_n)$, and the U_is are independent uniform(0,1) random variables. Since $Z_n \to 1$ in probability, $X_n \to Y$ in probability. Therefore, since the X_ns are bounded, X_n and Y are becoming perfectly correlated. In the top series of pictures on the right, you can see this case. Each graph has one hundred points plotted. Each point represents an instance of X_n and Y: its height is the value of X_n and its horizontal coordinate is the value of Y. You can easily see a linear relationship developing between X_n and Y.



For the theorem we required the random variables to be bounded. Is this really necessary? Yes. Consider the lower series of plots on the right. Here Y is again uniform(0,1). But X_n is defined as $X_n = Y + n \cdot Z_n$, where $Z_n = 1$ with probability $1/n$ and 0 otherwise. Thus, as n increases, fewer and fewer X_ns are far from Y (so $X_n \to Y$) but the distance from Y to those few X_ns is increasing. Here the correlation does not converge to 1.



Proof of main result

There are several proofs of this result. For brevity, I present a proof given by Prof. Luc Devroye, as it is shorter than my own proof.

Assume that X_n and Y are independent for all n. Since Y is nondegenerate, we can pick some $d > 0$ and some a such that $A^- = (-\infty, a]$ has probability $p > 0$ and $A^+ = (a + d, \infty)$ has probability $q > 0$.

Let $P(X_n \in A^-) = p_n$ and $P(X_n \in A^+) = q_n$. Since $X_n \to Y$ in probability, also $X_n \to Y$ in distribution. Thus, $p_n \to p$ and $q_n \to q$ as $n \to \infty$. Then $P(|X_n - Y| > d) \geq p_n q + p q_n$ by independence. But $p_n q + p q_n \to 2pq > 0$ as $n \to \infty$. Thus, for large enough n, $P(|X_n - Y| > d)$ is bounded below above 0, a contradiction to $X_n \to Y$ in probability. QED.