Comment aborder et distinguer les deux types de convergence : en probabilité et presque sûre ?

Définitions

La *convergence en probabilité* d'une suite de variables aléatoires X_n vers une variable X, toutes définies sur le même espace probabilisé $(\Omega;\ \mathcal{T};\ P)$, s'écrit :

$$\forall \eta > 0,\ \lim_{n\to+\infty} P\left(\{\omega\in\Omega/\ |X_n(\omega)-X(\omega)|\geq\eta\}\right)=0$$

La *convergence presque sûre* d'une suite de variables aléatoires X_n vers une variable X toutes définies sur le même espace probabilisé $(\Omega;\ \mathcal{T};\ P)$, s'écrit :

$$P\left(\left\{\omega\in\Omega/\ \lim_{n\to+\infty} X_n(\omega)=X(\omega)\right\}\right)=1$$

La convergence presque sûre entraîne la convergence en probabilité

Soit $\eta>0$. Dire que X_n converge en probabilité vers X signifie que, pour tout $\varepsilon>0$, il existe $N\in\mathbb{N}$ tel que pour tout $n\geq N$, $P\left(\{\omega\in\Omega/\ |X_n(\omega)-X(\omega)|\geq\eta\}\right)<\varepsilon$.
Dire que X_n converge presque sûrement vers X signifie que l'ensemble
$B_\eta=\bigcup_{n\in\mathbb{N}}\bigcap_{k\geq n}\{\omega/,|X_k(\omega)-X(\omega)|<\eta\}$ est de *probabilité un* pour tout $\eta>0$.
Donc, par passage au complémentaire, si on suppose la convergence presque sûre vers X, pour tout $\eta>0$, alors l'ensemble
$\tilde{B}_\eta=\bigcap_{n\in\mathbb{N}}\bigcup_{k\geq n}\{\omega/,|X_k(\omega)-X(\omega)|\geq\eta\}$ est de probabilité nulle.
Comme la suite $\bigcup_{k\geq n}\{\omega/,|X_k(\omega)-X(\omega)|\geq\eta\}$ est décroissante, on en déduit, grâce au théorème de convergence monotone, qu'à partir d'un certain rang N, $P\left(\bigcup_{k\geq N}\{\omega/,|X_k(\omega)-X(\omega)|\geq\eta\}\right)<\varepsilon$.
Or, pour tout $n\geq N$, $\{\omega\in\Omega/\ |X_n(\omega)-X(\omega)|\geq\eta\}\subset\bigcup_{k\geq N}\{\omega/,|X_k(\omega)-X(\omega)|\geq\eta\}$ ce qui prouve la convergence en probabilité.

Quelle est la différence principale entre convergence en probabilité et presque sûre ?

Si on reprend ce que l'on vient de faire, la convergence presque sûre entraîne que, pour tout $\eta>0$, la suite décroissante $C_\eta^{(n)}=\bigcup_{k\geq n}\{\omega/,|X_k(\omega)-X(\omega)|\geq\eta\}$ tend vers un ensemble de probabilité nulle.
Tandis que, pour la convergence en probabilité, ce sont les ensembles $\{\omega/\ |X_n(\omega)-X(\omega)|\geq\eta\}$ qui sont de probabilité aussi petite que l'on souhaite à partir d'un certain rang. Mais on ne sait rien sur leur limite supérieure, c'est à dire sur la limite de la suite $C_\eta^{(n)}$.
Pour obtenir un résultat sur cette limite supérieure, il faut procéder à une extraction et utiliser le théorème de Borel-Cantelli, qui n'est pas spécifiquement à votre programme.
On peut avec ces techniques, prouver que, de toute suite de variables qui converge en probabilité, on peut extraire une sous-suite qui converge presque sûrement.

Un contre exemple qui illustre cette différence

Je reprends un contre exemple traité en exercice.
On considère une suite de variables X_n indépendantes telle que, pour tout $n>0$, $P(X_n=0)=\dfrac{n-1}{n}$ et $P(X_n=1)=\dfrac{1}{n}$.

Il est clair que, pour tout $\eta > 0$, $P(|X_n| \geq \eta) \leq P(X_n = 1) = \dfrac{1}{n}$ et donc $\lim\limits_{n \to +\infty} P(|X_n| \geq \eta) = 0$, ce qui prouve la convergence en probabilité des X_n vers la variable aléatoire 0.
Montrons maintenant que les X_n ne convergent pas pour autant vers 0 presque sûrement. En effet, si tel était le cas, pour $\eta = \dfrac{1}{2}$, il existerait un rang N, tel que $P\left(\bigcap\limits_{n \geq N} \{\omega / \ |X_n| \leq \frac{1}{2}\}\right) \geq \frac{1}{2}$.
Or, pour tout entier M, compte-tenu de l'indépendance des X_n,
$P\left(\bigcap\limits_{N+M \geq n \geq N} \{\omega / \ |X_n| \leq \frac{1}{2}\}\right) = P\left(\bigcap\limits_{N+M \geq n \geq N} \{\omega / \ X_n = 0\}\right) = \prod\limits_{N+M \geq n \geq N} \dfrac{n-1}{n} = \dfrac{N-1}{N+M}$, car le produit est télescopique.
Par le théorème de convergence monotone, on en déduit :
$P\left(\bigcap\limits_{n \geq N} \{\omega / \ |X_n| \leq \frac{1}{2}\}\right) = \lim\limits_{M \to +\infty} \dfrac{N-1}{N+M} = 0$, ce qui est contradictoire avec la convergence presque sûre.
On remarque ici que ce qui pose problème, c'est bien la limite supérieure, dont la probabilité ne peut être égale à 1.