RÉSUMÉ DE COURS : CONVERGENCES ET APPROXIMATIONS

Dans tout ce qui suit, toutes les variables aléatoires en question seront supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

1 Convergence en probabilité

Définition 1.1. *Soit* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires et soit* X *une variable aléatoire. On dit que la suite* $(X_n)_{n\in\mathbb{N}^*}$ *converge en probabilité vers* X*, et on note* $X_n \xrightarrow{P} X$*, si pour tout réel* $\varepsilon > 0$*, on a :*

$$\lim_{n\to+\infty} P([|X_n - X| \geq \varepsilon]) = 0.$$

En d'autres termes, la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers X si la probabilité que X_n soit éloignée de X (d'au moins un écart $\varepsilon > 0$) tend vers 0 quand n tend vers $+\infty$. Grosso modo, cela signifie qu'il y a peu de chances que X_n soit éloignée de X pour n assez grand.

Théorème 1.2. (Inégalité de Markov) *Soit* X *une variable aléatoire (discrète ou à densité)* <u>*positive*</u> *admettant une espérance. Alors, pour tout réel* $\varepsilon > 0$ *:*

$$P([X \geq \varepsilon]) \leq \frac{E(X)}{\varepsilon}.$$

Théorème 1.3. (Inégalité de Bienaymé-Tchebychev) *Soit* X *une variable aléatoire (discrète ou à densité) admettant une variance. Alors, pour tout réel* $\varepsilon > 0$ *:*

$$P([|X - E(X)| \geq \varepsilon]) \leq \frac{V(X)}{\varepsilon^2}.$$

A noter que l'inégalité de Bienaymé-Tchebychev est une simple application de l'inégalité de Markov à la variable aléatoire $Y = (X - E(X))^2$. A noter aussi que l'inégalité de Bienaymé-Tchebychev est universelle au sens où elle ne dépend pas de la loi de X, mais uniquement de son espérance et de sa variance. Par contre, elle fournit un majorant assez grossier de $P([|X - E(X)| \geq \varepsilon])$. A noter enfin que l'inégalité de Bienaymé-Tchebychev permet de majorer la probabilité qu'une variable aléatoire X s'écarte de son espérance. Elle sert notamment à établir des propriétés de convergence en probabilité, comme par exemple le :

Théorème 1.4. (Loi faible des grands nombres) *Soit* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires (discrètes ou à densité)* <u>*indépendantes*</u> *admettant une même espérance* m *et une même variance, et posons* $\overline{X_n} = \frac{X_1+...+X_n}{n}$ *pour tout* $n \in \mathbb{N}^*$*. Alors la suite* $(\overline{X_n})$ *converge en probabilité vers la variable certaine égale à* m*, c'est-à-dire :*

$$\forall\varepsilon > 0, \quad \lim_{n\to+\infty} P([|\overline{X_n} - m| \geq \varepsilon]) = 0.$$

A noter que la loi faible des grands nombres s'obtient en appliquant l'inégalité de Bienaymé-Tchebychev à la variable aléatoire $\overline{X_n}$. En particulier, on voit que l'on peut remplacer l'hypothèse "variables (mutuellement) indépendantes" par "variables 2 à 2 indépendantes", voire par "variables non corrélées" et aboutir à la même conclusion. En d'autres termes, la loi faible des grands nombres signifie que la moyenne empirique $\overline{X_n}$ calculée sur une succession d'expériences aléatoires (dont les résultats X_k sont indépendants et ont une même espérance et une même variance) converge en probabilité vers l'espérance commune des X_k. D'un point de vue pratique, si l'on répète un grand nombre de fois et de manière indépendante une même expérience aléatoire, alors il y a de grandes chances que la moyenne empirique des résultats obtenus soit proche de l'espérance du résultat en général, ce qui justifie le fait a posteriori que l'on utilise des moyennes statistiques pour calculer l'espérance d'une variable aléatoire (avec une bonne approximation). Un cas particulier de la loi faible des grands nombres est donné par le :

Théorème 1.5. (Théorème d'or de Bernoulli) *Soit* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires indépendantes qui suivent toutes la loi de Bernoulli de paramètre* p*, et posons* $\overline{X_n} = \frac{X_1+...+X_n}{n}$ *pour tout* $n \in \mathbb{N}^*$*. Alors la suite* $(\overline{X_n})$ *converge en probabilité vers la variable certaine égale à* p*. Plus précisément, pour tout réel* $\varepsilon > 0$ *et tout entier* $n \in \mathbb{N}^*$*, on a :*

$$P([|\overline{X_n} - p| \geq \varepsilon]) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

En d'autres termes, le théorème d'or de Bernoulli signifie que, lors d'une succession d'épreuves de Bernoulli indépendantes et de même probabilité p, la fréquence statistique $\overline{X_n}$ de succès converge en probabilité vers la probabilité (théorique) p de succès. A noter que la convergence en probabilité est stable pour l'addition. Plus précisément, on a le :

Théorème 1.6. *Soient $(X_n)_{n\in\mathbb{N}^*}$ et $(Y_n)_{n\in\mathbb{N}^*}$ deux suites de variables aléatoires et soient X et Y deux variables aléatoires. Si $X_n \xrightarrow{P} X$ et si $Y_n \xrightarrow{P} Y$, alors $X_n + Y_n \xrightarrow{P} X + Y$.*

Enfin, on montre (et nous admettrons) le résultat suivant :

Théorème 1.7. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires qui converge en probabilité vers une variable aléatoire X, et soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue. Alors $f(X_n) \xrightarrow{P} f(X)$.*

A noter que, pour tout réel λ, la fonction $x \longmapsto \lambda x$ est continue sur $\mathbb{R}$. En particulier, on voit avec le théorème 1.7 que, si $X_n \xrightarrow{P} X$, alors $\lambda X_n \xrightarrow{P} \lambda X$. D'après le théorème 1.6, il s'ensuit que la convergence en probabilité est stable par combinaisons linéaires. Plus précisément, si $X_n \xrightarrow{P} X$ et si $Y_n \xrightarrow{P} Y$, alors $\lambda X_n + \mu Y_n \xrightarrow{P} \lambda X + \mu Y$ pour tous réels λ, μ. Comme ce résultat n'est pas officiellement au programme, il faut être en mesure de le retrouver avec les théorèmes 1.6 et 1.7 à chaque fois qu'on en aura besoin.

2 Convergence en loi

Définition 2.1. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires et soit X une variable aléatoire. On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X, et on note $X_n \xrightarrow{\mathcal{L}} X$, si pour tout réel x en lequel F_X est continue, on a :*

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

Grosso modo, cela signifie que la suite (X_n) converge en loi vers X si la loi de X_n (plus précisément sa fonction de répartition) converge point par point vers la loi de X (plus précisément sa fonction de répartition). A noter que l'hypothèse de continuité est automatiquement vérifiée si les X_n et X sont toutes des variables à densité. Si les variables aléatoires en question sont discrètes, on dispose du :

Théorème 2.2. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires <u>à valeurs dans $\mathbb{N}$</u> et soit X une variable aléatoire <u>à valeurs dans $\mathbb{N}$</u>. Alors la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X si et seulement si :*

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P([X_n = k]) = P([X = k]).$$

Un premier exemple de convergence en loi est donné par le :

Théorème 2.3. *Soit $\lambda \in \mathbb{R}_+^*$, et soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que $X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ pour tout $n \in \mathbb{N}^*$. Alors $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X telle que $X \hookrightarrow \mathcal{P}(\lambda)$.*

De façon générale, on peut montrer que *la convergence en probabilité implique la convergence en loi* $(*)$. Le résultat $(*)$ étant hors-programme, on ne pourra pas l'utiliser tel quel, mais le fait de le savoir permet dans bien des cas de deviner la limite en probabilité d'une suite de variables aléatoires à partir de la limite en loi. A noter que la réciproque de $(*)$ n'est pas vraie. Plus précisément, *il existe des suites de variables aléatoires qui convergent en loi mais pas en probabilité*. A noter aussi qu'on montre (et nous admettrons) le résultat suivant :

Théorème 2.4. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires qui converge en loi vers une variable aléatoire X, et soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue. Alors $f(X_n) \xrightarrow{L} f(X)$.*

Le cas le plus notable de convergence en loi est donné par le résultat suivant (qui est aussi admis) :

Théorème 2.5. (Théorème limite central) *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires (discrètes ou à densité) <u>indépendantes</u> et admettant toutes une espérance m et une variance $\sigma^2 \neq 0$, et posons pour tout $n \in \mathbb{N}^*$:*

$$\overline{X_n} = \frac{X_1 + ... + X_n}{n} \quad \text{et} \quad \overline{X_n}^* = \sqrt{n}\left(\frac{\overline{X_n} - m}{\sigma}\right).$$

Alors la suite $(\overline{X_n}^*)$ converge en loi vers une variable normale centrée réduite. Plus précisément, pour tous réels a, b avec $a < b$, on a :

$$\lim_{n\to+\infty} P([a \leq \overline{X_n}^* \leq b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}}\, dt.$$

A noter qu'ici, la variable aléatoire $\overline{X_n}^*$ n'est ni plus ni moins que la variable aléatoire centrée réduite associée à $\overline{X_n}$, puisque $E(\overline{X_n}) = m$ et $V(\overline{X_n}) = \frac{\sigma^2}{n}$ pour tout $n \in \mathbb{N}^*$. A noter aussi le caractère universel du théorème limite central, puisque la suite $(\overline{X_n}^*)$ converge en loi vers une variable aléatoire dont la loi est indépendante de la loi commune des X_k.

3 Approximations

3.1 Approximation de la loi binomiale par la loi normale

Considérons une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes qui suivent toutes une loi de Bernoulli de même paramètre p. Alors on sait d'après la propriété de stabilité par addition de la loi binomiale que la variable aléatoire $S_n = X_1 + ... + X_n$ suit la loi binomiale de paramètres n, p. De plus, le théorème limite central permet d'affirmer que la suite (S_n^*) converge en loi vers une variable normale centrée réduite (vu que $S_n^* = \overline{X_n}^*$ pour tout $n \in \mathbb{N}^*$). En particulier, pour n assez grand, la variable aléatoire S_n est bien approchée par une variable aléatoire suivant la loi normale d'espérance np et de variance npq. Dans la pratique, on approche une loi binomiale par une loi normale si n est assez grand et si p n'est ni trop proche de 0, ni de 1. Plus précisément :

$$\boxed{\text{on approche la loi binomiale } \mathcal{B}(n,p) \text{ par la loi normale } \mathcal{N}(np, npq) \text{ si } n \geq 20 \text{ et } p \simeq 0,5.}$$

A noter que les conditions d'approximation diffèrent suivant les auteurs mais que, dans tous les cas et conformément au programme officiel : "*toutes les indications devront être fournies aux candidats quant à la justification de l'utilisation des approximations*".

3.2 Approximation de la loi de Poisson par la loi normale

Soit μ un réel > 0, et considérons une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes qui suivent toutes la loi de Poisson $\mathcal{P}(\mu)$. Alors on sait d'après la propriété de stabilité par addition de la loi de Poisson que la variable aléatoire $S_n = X_1 + ... + X_n$ suit la loi de Poisson $\mathcal{P}(n\mu)$. De plus, le théorème limite central permet d'affirmer que la suite (S_n^*) converge en loi vers une variable normale centrée réduite (vu que, comme précédemment, on a $S_n^* = \overline{X_n}^*$ pour tout $n \in \mathbb{N}^*$). En particulier, pour n assez grand, la variable aléatoire S_n est bien approchée par une variable aléatoire suivant la loi normale d'espérance $n\mu$ et de variance $n\mu$. Dans la pratique, on approche une loi de Poisson de paramètre λ par une loi normale d'espérance λ et de variance λ si λ est assez grand. Plus précisément :

$$\boxed{\text{on approche la loi de Poisson } \mathcal{P}(\lambda) \text{ par la loi normale } \mathcal{N}(\lambda, \lambda) \text{ si } \lambda \geq 10.}$$