BCPST2 9 5 2

14 Théorèmes limite

I Rappel et définition

A) Variable centrée réduite

Définition :

Soit X une variable aléatoire réelle admettant une espérance m et une variance σ^2 non nulle.
La variable $X^* = \dfrac{X-m}{\sigma}$ est appellée variable centrée et réduite associée à X. Elle vérifie :

$$\mathbf{E}(X^*) = 0, \quad \sigma(X^*) = 1$$

Démonstration :

B) Moyenne de variables aléatoire

Proposition :

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes admettant la même espérance m et la même variance σ^2 non nulle.
Alors $M_n = \dfrac{1}{n}\sum_{k=1}^{n} X_k$ vérifie :

$$\mathbf{E}(M_n) = m, \quad \sigma(M_n) = \frac{\sigma}{\sqrt{n}}$$

Démonstration :

II Loi faible des grands nombres

Théorème : Loi faible des grands nombres

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de même espérance m et de même écart-type σ.
Posons, pour $n \in \mathbb{N}^*$

$$M_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

Alors :

$$\forall \varepsilon > 0, \ P\left([|M_n - m| \geqslant \varepsilon]\right) \underset{n\to+\infty}{\longrightarrow} 0$$

On dit que (M_n) converge en probabilité vers la loi certaine m

Démonstration : Appliquons l'inégalité de Bienaymé-Tchebychev à M_n.

Exemple : Fréquence d'apparition du 6 dans un tir de dé :

♦
- ✧ Quelle est la limite en probabilité de la fréquence d'apparition d'un 6 dans des tirs de dés successifs ?
- ✧ Majorer la probabilité d'avoir tiré moins d'une fois sur 10 le chiffre 6, sur 10 tirs, sur 1000 tirs ou sur 10000 tirs.

Remarque: Loi forte des grands nombres

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indentiquement distribuées et admettant une espérance. Alors,
$$\frac{X_1+\cdots+X_n}{n} \underset{n\to+\infty}{\longrightarrow} \mathbf{E}(X_n) \text{ p.s.}$$
(C'est à dire que c'est un évènement de probabilité 1)
En particulier, cela justifie les calculs expérimentaux d'espérance et de loi par simulations de variables aléatoires.

On suppose qu'on dispose d'une fonction `simule()` qui simule une variable aléatoire.

Calcul expérimental d'une espérance

```
def esperance():
        S = 0
        for k in range(1000):
                S +=simule()
        return S/1000
```

III Convergence en loi dans le cas discret

A) Définition

Définition :

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires *discrètes à valeurs dans* $\mathbb{N}$, et X une variable aléatoire *discrète à valeurs dans* $\mathbb{N}$.
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ *converge en loi* vers X si et seulement si
$$\forall k\in\mathbb{N} \qquad P\left([X_n=k]\right) \underset{n\to+\infty}{\longrightarrow} P([X=k])$$

Remarque:

Le terme «convergence en loi» n'est pas au programme. Cependant, les résultats suivants, exprimés en terme de limite, le sont.

B) Cas de la loi hypergéométrique

Théorème :

Soit $n\in\mathbb{N}$ et $p\in\mathbb{Q}$.
Soit $(X_N)_{N\in\mathbb{N}}$ une suite de variables aléatoires telle que
$$X_N \hookrightarrow \mathcal{H}(N,n,p)$$
On rappelle qu'on a donc $Np\in\mathbb{N}$, la suite est en fait définie que pour certaines valeurs de N.
Alors pour tout $k\in[\![0,n]\!]$
$$P\left([X_N=k]\right) \underset{N\to+\infty}{\longrightarrow} \binom{n}{k}p^k\left(1-p\right)^{n-k}$$
Autrement dit $(X_N)_{N\in\mathbb{N}}$ converge en loi vers toute variable aléatoire suivant la loi binomiale de paramètres n et p.

Démonstration :

Remarque:

En pratique, quand $N\geqslant 10n$, on considère qu'on peut approcher la loi $\mathcal{H}(N,n,p)$ par la loi $\mathcal{B}(n,p)$.

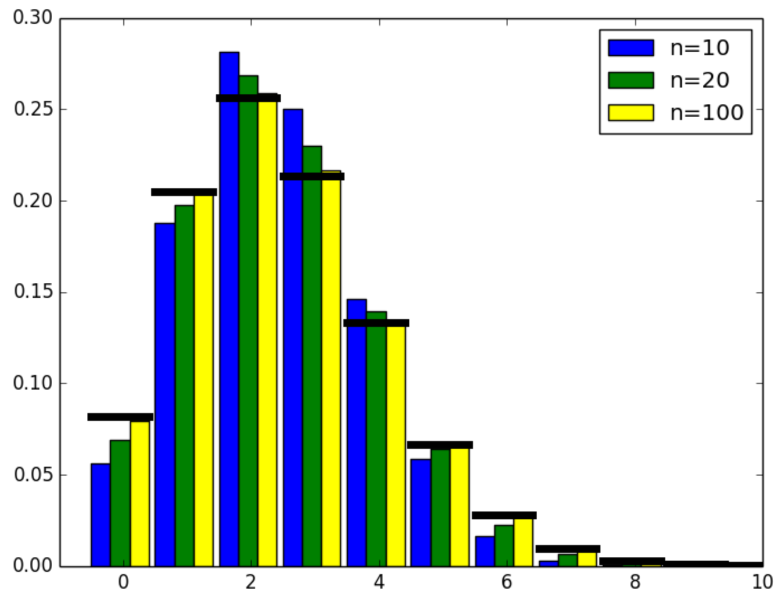
C) Cas de la loi Binomiale

Théorème :

Soit $\lambda > 0$ et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour tout $n \in \mathbb{N}$, $X_n \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})$.
Alors pour tout $k \in [\![0, n]\!]$

$$P\left([X_n = k]\right) \underset{n\to+\infty}{\longrightarrow} \frac{\lambda^k}{k!}e^{-\lambda}$$

Autrement dit $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable suivant la loi de Poisson de paramètre λ.



Remarque:

En pratique, on approchera une loi binomiale de paramètres n et p par une loi de Poisson de paramètre np dès que $p \leqslant 0,1$ et $n \geqslant 30$.

IV Convergence en loi dans le cas général : Théorème central limite

A) Convergence en loi

Définition :

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles et X une variable aléatoire ; notons F la fonction de répartition de X. On suppose ici que F est continue.
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ *converge en loi* vers X si et seulement si

$$\forall x \in I \qquad F_n(x) \underset{n\to+\infty}{\longrightarrow} F(x)$$

Remarque:

✧ Dans ce cas, on a aussi :

$$\forall a, b \in \mathbb{R}, P(a < X_n < b) \underset{n\to+\infty}{\longrightarrow} P(a < X < b)$$

✧ On admettra que cette défintion prolonge la définition donnée dans le cas discret.

Remarque:

Comme précédement le terme «convergence en loi» n'est pas au programme, mais les résultats suivants, exprimés en termes de limite le sont.

B) Théorème central limite

Le théorème central-limite, est un résultat remarquable : avec des hypothèses extrêmement réduites (une suite de variables aléatoires indépendantes de même loi admettant une variance non-nulle), il affirme que la suite des variables centrées réduites associées aux sommes partielles de cette suite converge en loi vers la loi normale centrée réduite.

Théorème : central limite

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, *de même loi*, admettant une variance non-nulle.
Notons μ l'espérance commune des X_n pour $n \in \mathbb{N}^*$ et σ leur écart-type commun.
Posons alors, pour $n \in \mathbb{N}^*$,

$$M_n = \frac{1}{n}\sum_{k=1}^{n} X_k \text{ et } M_n^* = \frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Alors $(M_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.
Autrement dit, en notant Φ la fonction de répartition la loi $\mathcal{N}(0,1)$, on a, pour tous réels a et b vérifiant $a < b$

$$\boxed{P\left(a < \frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}} < b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a) = \int_a^b \frac{\exp(-\frac{x^2}{2})}{\sqrt{2\pi}}\,\mathrm{d}x}$$

Remarque:

On écrit aussi le même théorème sous la forme suivante :
On pose $S_n = \displaystyle\sum_{k=1}^{n} X_k$ et $S_n^* = \dfrac{S_n - n\mu}{\sqrt{n}\sigma}$.

$$\boxed{P\left(a < \frac{S_n - n\mu}{\sqrt{n}\sigma} < b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a)}$$

C) Cas de la loi Binomiale

Théorème : de Moivre-Laplace

Soit $p \in]0,1[$ et $(S_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que

$$\forall n \in \mathbb{N}^*, S_n \hookrightarrow \mathcal{B}(n,p)$$

Alors :

$$\boxed{P\left(a \leq S_n^* \leq b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a)}$$

ou encore :

$$\boxed{P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a)}$$

Démonstration :

Mise en place de l'approximation

Méthode

Soit $p \in]0,1[$ et $(S_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telles que $S_n \hookrightarrow \mathcal{B}(n,p)$.

On veut avoir une approximation de $P\left(a \leq S_n \leq b\right)$:

$$\begin{aligned} P(a \leqslant S_n \leqslant b) &= P\left(\tfrac{a-np}{\sqrt{npq}} \leqslant S_n^* \leqslant \tfrac{b-np}{\sqrt{npq}}\right) \\ &\approx \Phi\left(\tfrac{b-np}{\sqrt{npq}}\right) - \Phi\left(\tfrac{a-np}{\sqrt{npq}}\right) \end{aligned}$$

Autrement dit, on approche la loi binomiale $\mathcal{B}(n,p)$ par la loi normale $\mathcal{N}(np, \sqrt{npq})$.

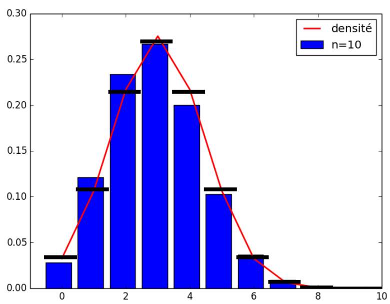
Les conditions $n \geqslant 30$, $np \geqslant 5$ et $nq \geqslant 5$ sont généralement considérées comme suffisantes.

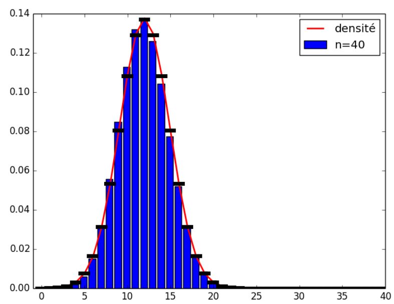
Remarque:

Si on veut obtenir une approximation de $P(S_n = k)$, prendre $a = b = k$ n'est pas raisonnable puisque l'on obtiendrait alors 0.

Pour obtenir une approximation raisonnable, on applique une *correction de continuité* : pour tout entier k, on approche la valeur de $P(S_n = k)$ par $P\left(k - \frac{1}{2} \leqslant S_n \leqslant k + \frac{1}{2}\right)$ ce qui donne les approximations :

$$\boxed{P(S_n = k) \approx \Phi\left(\frac{k + \frac{1}{2} - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k - \frac{1}{2} - np}{\sqrt{npq}}\right)}$$





Exemple :

♦ On lance 50 fois une pièce équilibrée. Quelle est la probabilité que le nombre de piles obtenus soit compris entre 20 et 30 ?

D) Cas de la loi de Poisson

Théorème :

Soit $\lambda > 0$ et soit $(S_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires vérifiant

$$\forall n \in \mathbb{N}^* \qquad S_n \hookrightarrow \mathcal{P}(n\lambda)$$

Alors :

$$\boxed{\mathbf{P}\left(a \leq S_n^* \leq b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a)}$$

ou encore :

$$\boxed{\mathbf{P}\left(a \leq \frac{S_n - n\lambda}{\sqrt{n\lambda}} \leq b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a)}$$

Démonstration :

Mise en place de l'approximation

Méthode

Soit $(S_n)_{n\geq 0}$ une suite de variables aléatoire suivant une loi $\mathcal{P}(n\lambda)$.
On veut avoir une approximation de $P\left(a \leq S_n \leq b\right)$:

$$\begin{aligned}\mathbf{P}(a \leqslant S_n \leqslant b) &= \mathbf{P}\left(\tfrac{a-n\lambda}{\sqrt{n\lambda}} \leqslant S_n^* \leqslant \tfrac{b-n\lambda}{\sqrt{n\lambda}}\right)\\ &\approx \Phi\left(\tfrac{b-n\lambda}{\sqrt{n\lambda}}\right) - \Phi\left(\tfrac{a-n\lambda}{\sqrt{n\lambda}}\right)\end{aligned}$$

Autrement dit, on approche la loi de poisson $\mathcal{P}(\mu)$ par la loi normale $\mathcal{N}(\mu, \sqrt{\mu})$.

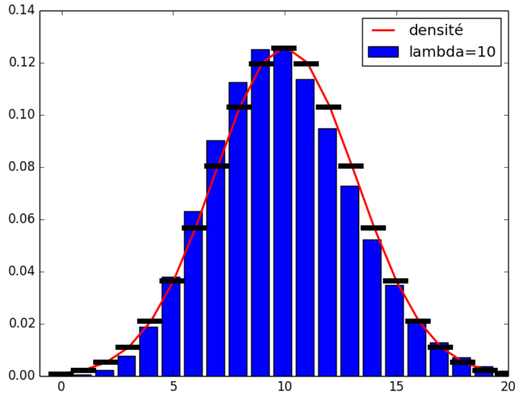
La condition $\mu \geq 15$ est généralement considérée comme suffisante.

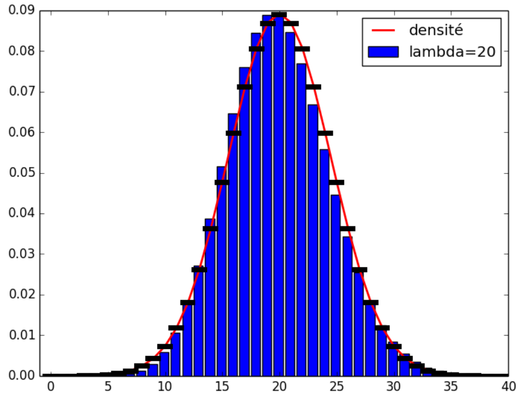
Remarque:

Si on veut obtenir une approximation de $P(S_n = k)$, prendre $a = b = k$ n'est pas raisonnable puisque l'on obtiendrait alors 0.

Pour obtenir une approximation raisonnable, on applique une *correction de continuité* : pour tout entier k, on approche la valeur de $\mathbf{P}(S_n = k)$ par $\mathbf{P}\left(k - \frac{1}{2} \leqslant S_n \leqslant k + \frac{1}{2}\right)$ ce qui donne les approximations :

$$\boxed{\mathbf{P}(S_n = k) \approx \Phi\left(\frac{k + \frac{1}{2} - n\lambda}{\sqrt{n\lambda}}\right) - \Phi\left(\frac{k - \frac{1}{2} - n\lambda}{\sqrt{n\lambda}}\right)}$$





Théorème : Théorème central limite deuxième forme

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, *de même loi*, admettant une variance non-nulle.
Notons μ l'espérance commune des X_n pour $n \in \mathbb{N}^*$ et σ leur écart-type commun. Posons alors, pour $n \in \mathbb{N}^*$,

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

On note de plus :

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{et } \sigma' = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(X_k - \overline{X_n}\right)^2}$$

$$\boxed{P\left(a < \frac{Y_n - \mu}{\frac{\sigma'}{\sqrt{n}}} < b\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(a) = \int_a^b \frac{\exp(-\frac{x^2}{2})}{\sqrt{2\pi}}\,\mathrm{d}x}$$

Remarque:

L'intérêt de cette deuxième forme provient du fait que lorsqu'on fait une analyse statistique, on ne connait pas nécessairement à l'avance l'écart-type σ. On le calcule donc au fur et à mesure des expériences.

V Résumé

Loi	Loi approchée	Conditions
$\mathcal{H}(N,n,p)$	$\mathcal{B}(n,p)$	$N \geqslant 10n$
$\mathcal{B}(n,p)$	$\mathcal{P}(np)$	$p \leqslant 0,1$ et $n \geqslant 30$
$\mathcal{B}(n,p)$	$\mathcal{N}(np, \sqrt{npq})$	$n \geqslant 30,\ np \geqslant 5$ et $nq \geqslant 5$
$\mathcal{P}(\mu)$	$\mathcal{N}(\mu, \sqrt{\mu})$	$\mu \geqslant 15$

BCPST2 952

14 Théorèmes limites

«Quand les lois mathématiques s'appliquent à la réalité, elles ne sont pas exactes, et quand elles sont exactes, elles ne s'appliquent pas à la réalité.» Albert Einstein

♦ Exercice 1:

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variable aléatoire indépendantes suivant toutes une loi de Bernoulli de paramètre p.

Pour tout entier $n \in \mathbb{N}^*$, on pose $M_n = \dfrac{Y_1 + \cdots + Y_n}{n}$ et $Y_n = X_n + X_{n+1}$

1°) Rappeler la loi faible des grands nombres.

2°) Les variables aléatoires Y_n sont-elles indépendantes ?

3°) Calculer l'espérance et la variance de M_n .

4°) À l'aide de l'inégalité de Bienaymé-Tchebychev, démontrer que, pour tout $\varepsilon > 0$, on a $\lim\limits_{n\to+\infty} \mathbf{P}(|M_n - 2p| > \varepsilon) = 0$.

♦ Exercice 2:

On reproduit une expérience de probabilité de succès p.

On pose $S_0 = 0$. Pour tout $n \in \mathbb{N}^*$, on note S_n le nombre d'échecs avant le n-ième succès, et $X_n = S_n - S_{n-1}$.

1°) Reconnaître la loi de X_n et donner son espérance et sa variance.

En déduire l'espérance et la variance de S_n.

2°) Montrer, pour tout $k \in \mathbb{N}$ et pour tout $n \in \mathbb{N}^*$

$$\mathbf{P}(S_n = k) = \binom{n+k-1}{k} p^n (1-p)^k$$

3°) Montrer

$$\lim_{n\to+\infty} \sum_{k=0}^{\frac{n(1-p)}{p}} \binom{n+k-1}{k} p^n (1-p)^k = \frac{1}{2}$$

♦ Exercice 3:

On se donne une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ telle que X_1 suit la loi uniforme sur $[0,1]$ et que pour tout $n \geqslant 2$ et tout $x \in]0,1]$, la loi conditionnelle de X_n sachant $X_{n-1} = x$ est la loi uniforme sur $[0, x]$.

1°) Déterminer une densité de X_2.

2°) Montrer que pour tout $n \geqslant 2$, l'application

$$f_n : x \mapsto \begin{cases} \frac{(-\ln x)^{n-1}}{(n-1)!} & \text{si } x \in]0,1] \\ 0 & \text{sinon} \end{cases}$$

est une densité de X_n.

3°) Montrer que pour tout $n \in \mathbb{N}^*$, X_n admet une espérance et une variance.
Exprimer $\mathbf{E}(X_{n+1})$ en fonction de $\mathbf{E}(X_n)$ et en déduire une expression de $\mathbf{E}(X_n)$.
Exprimer $\mathbf{E}(X_{n+1}^2)$ en fonction de $\mathbf{E}(X_n^2)$ et en déduire une expression de $\mathbf{V}(X_n)$.

4°) **a)** Montrer, pour tout $n \geqslant 2$ et tout $t \in]0,1]$

$$\mathbf{P}(X_n \leqslant t) = \frac{t(-\ln t)^{n-1}}{(n-1)!} + \mathbf{P}(X_{n-1} \leqslant t)$$

b) En déduire que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable certaine qu'on précisera

♦ Exercice 4:

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables indépendantes suivant chacune la loi de Poisson de paramètre 1. On pose, pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{k=1}^{n} X_k$.

1°) Pour tout $n \in \mathbb{N}^*$, déterminer la loi de S_n puis calculer $\mathbf{P}(S_n \leqslant n)$.

2°) Montrer

$$\lim_{n\to+\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}$$

♦ Exercice 5:

Un chef de projet informatique estime la taille du projet qu'on lui demande de réaliser à environ 10000 fonctions. Chaque jour, chacun de ses 10 ingénieurs code un nombre de fonction. Chaque réalisation demande un temps aléatoire, suivant une certaine loi. La durée de codage d'une fonction suit une loi à densité, exprimée en journées. Estimer la durée du projet (en nombre de jours) dans les cas où la loi suivie par la durée du codage d'une fonction

1°) Suit la loi uniforme sur l'intervalle $[0, \frac{1}{2}]$.

2°) Suit la loi exponentielle de paramètre 4.
Pour donner un intervalle de confiance sur la durée du projet, on utilisera d'une part l'inégalité de Bienaymé-Tchebychev et d'autre part l'approximation par une loi normale et on comparera les résultats.

♦ Exercice 6:

Un étudiant fait en moyenne une faute d'orthographe tous les 500 mots.
Quelle est la probabilité qu'il ne fasse pas plus de 5 fautes dans un texte contenant 200 mots, dans un texte contenant 2000 mots ?

♦ Exercice 7:

1200 élèves déjeunent à la cantine du lycée, qui propose fromage ou dessert.

On s'attend à ce que chaque élève choisisse le dessert avec probabilité $0,75$. Combien faut-il prévoir de fromages et de desserts de façon à ce qu'il y ait 95% de chances que mêmes les derniers aient le choix ?

♦ Exercice 8:

Deux compagnies maritimes (A et B) affrètent chacune 1000 bateaux. Le coût de remplacement d'un bateau est de $200000E$. La probabilité de perdre un bateau est de $p = 0,001$. On note X (resp. Y) le nombre de bateaux perdus par A (resp. B).

1°) Quelles sont les lois suivies par X et Y ? Par quelles lois peut-on les approcher ?

2°) Quelle réserve financière doit faire chacune des compagnies pour être en mesure de remplacer ses bateaux perdus avec une probabilité supérieure ou égale à $0,999$?

3°) Serait-il intéressant pour les deux compagnies de fusionner ?