Chapitre 18 : Loi des grands nombres

1 Loi binomiale

Propriété

Toute variable aléatoire suivant une loi binomiale peut s'écrire comme une somme de variables aléatoires de Bernoulli indépendantes et identiquement distribuées.
Autrement dit : si $X \sim \mathcal{B}(n,p)$ alors $X = X_1 + X_2 + \ldots + X_n$ avec X_1, X_2, $\ldots$, X_n des variables aléatoires indépendantes suivant la loi $\mathcal{B}(p)$

Exemple

Si X suit une loi binomiale de paramètres $n = 3$ et $p = 0,4$. alors :
$X = X_1 + X_2 + X_3$ où X_1, X_2, X_3 des variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre $p = 0,4$.

Propriété

Si X suit une loi binomiale de paramètres n et p, alors :

1. $E(X) = \ldots$
2. $V(X) = \ldots\ldots$
3. $\sigma(x) = \ldots\ldots$

Démonstration exigible :

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p. Alors, il existe n variables aléatoires de Bernoulli de paramètres p telles que $X = X_1 + ... + X_n$.

Ainsi, pour tout $k \in \{1; ...; n\}$, $E(X_k) = \ldots\ldots$ et $V(X_k) = \ldots\ldots$.

• Or, $E(X) = E(X_1 + ... + X_n) =$ $\blacksquare$

• Les variables aléatoires $X_1, ...X_n$ étant indépendantes, par définition du schéma de Bernoulli, on a :

$V(X) =$

$\blacksquare$

• $\sigma(X) =$ $\blacksquare$

Exemple

Soit X une variable aléatoire suivant une loi binomiale de paramètres $n = 20$ et $p = 0,3$.

$E(X) =$ $\qquad$ $V(X) =$ $\qquad$ $\sigma(X) =$

2 Échantillons de n variables aléatoire identiques et indépendantes

Définition

Une liste $(X_1, X_2, \ldots, X_n)$ de variables aléatoires indépendantes suivants toutes la même loi est appelée **échantillon de taille n** associée à cette loi (ou à une variable aléatoire suivant cette loi).

Exemple

On prélève avec remise une pièce métallique dans une chaîne de production et on relève sa masse, en grammes.
Soit X la variable aléatoire qui associe la masse de la pièce métallique prélevée.
On note X_1 la variable aléatoire qui associe la masse, en grammes, de la pièce métallique obtenue lors du premier prélèvement, X_2 la variable aléatoire associant la masse de la pièce métallique du deuxième prélèvement et ainsi de suite.
La liste (X_1, X_2, X_3) est constituée de trois variables aléatoires identiques et indépendantes. Elles suivent la même loi : celle de X.
La liste (X_1, X_2, X_3) est un échantillon de taille 3 de la loi de X.

Définition

Soit une variable aléatoire X et soit un échantillon $(X_1, X_2, ..., X_n)$ de taille n de variables aléatoires indépendantes suivant une même loi que X.

- La **variable aléatoire somme** S_n de l'échantillon est donnée par : $S_n = X_1 + X_2 + ... + X_n$
- La **variable aléatoire moyenne** M_n de l'échantillon est donnée par : $M_n = \dfrac{1}{n}(X_1 + X_2 + ... + X_n)$

Exemple

Dans l'exemple précédent :

- La variable aléatoire S_3 associé aux trois tirages la somme de la masse des trois pièces métalliques
- La variable aléatoire M_3 associe la masse moyenne d'une pièce métallique d'un lot de trois pièces métalliques.

Propriété

Soit une variable aléatoire X et soit un échantillon $(X_1, X_2, ..., X_n)$ de taille n de variables aléatoires indépendantes suivant une même loi que X.

$$E(S_n) = \ldots\ldots \qquad V(S_n) = \ldots\ldots \qquad \sigma(S_n) = \ldots\ldots$$

$$E(M_n) = \ldots\ldots \qquad V(M_n) = \ldots\ldots \qquad \sigma(M_n) = \ldots\ldots$$

Démonstration :

• La linéarité de l'espérance donne $E(S_n) = E(X_1) + ... + E(X_n)$.
Or ces variables aléatoires suivent la même loi. Elles ont donc la même espérance. D'où, $E(S_n) =$ $\blacksquare$

• De même, les variables aléatoires $X_1, ...X_n$ étant indépendantes, on a : $V(S_n) =$ $\blacksquare$

• $\sigma(S_n) = \sqrt{V(S_n)} =$ $\blacksquare$

• La linéarité de l'espérance et la propriété précédente donnent

$$E(M_n) = E\left(\frac{X_1 + ... + X_n}{n}\right) =$$

$\blacksquare$

• On sait que, pour tout $a \in \mathbb{R}$, on a $V(aS_n) = a^2V(S_n)$
Donc $V(M_n) = V\left(\frac{S_n}{n}\right) =$ $\blacksquare$

• $\sigma(M_n) = \sqrt{V(M_n)} = \sqrt{\frac{V(X_k)}{n}} =$ $\blacksquare$

Exemple

On lance 5 dés équilibrés à six faces. on note X la variable aléatoire correspondant à la somme des résultats obtenus. Calculons $E(X)$ et $V(X)$.

Pour $k \in \{1; ...; 5\}$, on note X_k la variable aléatoire correspondant au résultat du dé numéro k.
On a alors $X = X_1 + X_2 + X_3 + X_4 + X_5$.
Chaque dé étant équilibré, toutes ces variables aléatoires suivent la même loi de probabilité.

1. Déterminons E(X) :
   On a $E(X) =$.
   Or la loi de probabilités de X_1 est :

x_i	1	2	3	4	5	6
$P(X_1 = x_i)$						

   On a $E(X_1) =$. Donc $E(X) =$
2. Déterminons V(X) :
   Les variables aléatoires $X_1; ...; X_n$ étant indépendantes et identiquement distribuées, on a $V(X) = 5V(X_1)$.
   Or $V(X_1) = \frac{1}{6} \times (1 - 3,5)^2 + ... + \frac{1}{6} \times (6 - 3,5)^2 = \frac{35}{12}$, d'où $V(X) = 5 \times \frac{35}{12} = \frac{175}{12}$.

3 Inégalité de Bienaymé-Tchebytchev

Propriété (Admise)

Soit une variable aléatoire X. Pour tout réel strictement positif δ, on a :

$$P\big(|X - E(X)| \geqslant \delta\big) \leqslant \frac{V(X)}{\delta^2}$$

Remarque

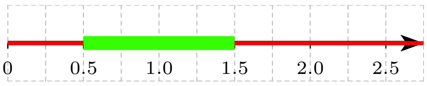
Cette inégalité signifie que la probabilité qu'une variable aléatoire X prenne une certaine valeur diminue d'autant plus que cette valeur est éloignée de l'espérance de X.

Exemple

Le taux moyen de glycémie d'une personne est de $1g.L^{-1}$ avec une variance de $0,1$.

Une personne présente une taux X critique si son taux ne se situe pas dans l'intervalle $]0.5; 1.5[$.

Estimer la probabilité qu'une personne présente un taux critique



Le taux X est critique $\Leftrightarrow$ ....

Ainsi, par l'Inégalité de Bienaymé-Tchebychev :

$P(|X - E(X)| \geq 0.5) \leq$ soit $P(|X - E(X)| \geq 0.5) \leq$.

La probabilité qu'une personne présente un taux critique est inférieure ou égale à 0.4.

Exemple

Soit une variable aléatoire X qui suit la loi binomiale de paramètres $n = 20$ et $p = 0,1$.

1. Appliquer l'inégalité de Bienaymé-Tchebytchev avec $\delta = 2\sigma(X)$. Interpréter.
2. Recommencer avec $\delta = 3\sigma(X)$ et $\delta = 4\sigma(X)$. Que constate-t-on ?

1. $E(X) =$ $V(X) =$ $\sigma(X) =$

   $P\left(|X - E(X)| \geqslant 2\sigma(X)\right) \leqslant$ donc $P\left(|X-2)| \geqslant 2\sqrt{1,8}\right) \leqslant$

   La probabilité que l'écart de X à $E(X)$ soit supérieur à $2\sigma(X)$ est majorée par $0,25$.

2.
   - pour $\delta = 3\sigma(X)$: $P\left(|X - E(X)| \geqslant 3\sigma(X)\right) \leqslant$ donc $P\left(|X-2)| \geqslant 3\sqrt{1,8}\right) \leqslant$
   - pour $\delta = 4\sigma(X)$: $P\left(|X - E(X)| \geqslant 4\sigma(X)\right) \leqslant$ donc $P\left(|X-2)| \geqslant 4\sqrt{1,8}\right) \leqslant$
   - On peut en déduire que les écarts de X à $E(X)$ de quelques σ deviennent improbables.

4 Inégalité de concentration

Propriété (Admise)

Soit une variable aléatoire moyenne M_n d'un échantillon de taille n de la variable aléatoire X. Pour tout réel strictement positif δ, on a :

$$P\big(|M_n - E(X)| \geqslant \delta\big) \leqslant \frac{V(X)}{n\delta^2}$$

Exemple

Soit une variable aléatoire X qui suit la loi de Bernoulli de paramètre $0,2$. On considère un échantillon de n variables aléatoires suivant la loi de X. On appelle M_n la variable aléatoire moyenne associée à cet échantillon. Déterminer la taille n de l'échantillon tel que la probabilité que la moyenne M_n appartienne à l'intervalle $]0,03; 0,37[$ soit supérieur à $0,95$.

On cherche à calculer n tel que $P(0,03 < M_n < 0,37) \geqslant 0,95$. Dans l'idée d'appliquer l'inégalité de concentration, on fait apparaitre l'espérance de X dans l'inégalité. Or $E(X) = p = 0,2$.
Ainsi, on cherche n tel que :

$$\begin{aligned} P(0,03 - 0,2 < M_n - 0,2 < 0,37 - 0,2) &\geqslant 0,95 \\ P(-0,17 < M_n - 0,2 < 0,17) &\geqslant 0,95 \\ P(|M_n - 0,2| < 0,17) &\geqslant 0,95 \\ 1 - P(|M_n - 0,2| > 0,17) &\geqslant 0,95 \\ P(|M_n - 0,2| > 0,17) &\leqslant 0,05 \end{aligned}$$

En prenant $\delta = 0,17$ dans l'inégalité de concentration, on a : $P\big(|M_n - E(X)| \geqslant \delta\big) \leqslant \dfrac{V(X)}{n\delta^2}$.
Or, $V(X) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

On cherche un entier n tel que : $\dfrac{0,16}{n0,17^2} \leqslant 0,05$

$$\begin{aligned} \frac{0,16}{n0,17^2} \leqslant 0,05 &\Longleftrightarrow \frac{n0.17^2}{0.16} \geq \frac{1}{0.05} \text{ (par inverse)} \\ \frac{0,16}{n0,17^2} \leqslant 0,05 &\Longleftrightarrow n \geq \frac{1}{0.05} \times \frac{0.16}{0.17^2} \\ \frac{0,16}{n0,17^2} \leqslant 0,05 &\Longleftrightarrow n \geq \ldots\ldots \end{aligned}$$

Pour $n \geqslant 111$, la probabilité que la moyenne M_n appartienne à l'intervalle $]0,03; 0,37[$ est supérieur à $0,95$.

5 Loi des grands nombres

Propriété (Admise)

Soit une variable aléatoire moyenne M_n d'un échantillon de taille n de la variable aléatoire X. Pour tout réel strictement positif δ, on a :

$$\lim_{n\to+\infty} P\big(|M_n - E(X)| \geqslant \delta\big) = \ldots$$

Remarque

La loi des grands nombres illustre le fait que la moyenne de l'échantillon se rapproche de l'espérance des variables aléatoires quand la taille de l'échantillon "devient grande"

Exemple

On considère un jeu de 52 cartes.

On tire treize cartes avec remise et on note le nombre de rois obtenus.

D'après la loi des grands nombres, si on répète cette expérience un grand nombre n de fois et que l'on note M_n la moyenne des résultats obtenus, vers quelle valeur M_n converge t-elle ?

D'après la loi des grands nombres, M_n va tendre vers l'espérance de cette variable aléatoire.

Ici la variable aléatoire X correspondant à cette expérience suit une loi binomiale de paramètres $n = \ldots\ldots$ et $p = \ldots\ldots$.

Donc la limite de M_n est $\ldots\ldots\ldots\ldots\ldots\ldots$.