Chapitre 13: Variables aléatoires et loi des grands nombres.

I. Somme de variables aléatoires

Exemple :
On considère deux jeux dont les gains sont donnés :
- pour le premier jeu, par la variable aléatoire X qui prend les valeurs 1 et 2.
- pour le second jeu, par la variable aléatoire Y qui prend les valeurs –2, 3 et 4.

Par exemple, l'évènement $(X = 1) \cap (Y = -2)$ signifie qu'on a gagné 1 € au premier jeu et perdu 2 € au deuxième jeu.
Considérons la variable aléatoire somme $X + Y$ donnant le gain total cumulé aux deux jeux.
Alors la variable aléatoire $X + Y$ peut prendre les valeurs : -1, 0, 4, 5 et 6.
En effet, on a par exemple $X + Y = 0$ avec $(X = 2) \cap (Y = -2)$.

Par ailleurs, pour calculer par exemple, la probabilité de l'évènement $X + Y = 5$, on pourra calculer : $P(X + Y = 5) = P\big((X = 1) \cap (Y = 4)\big) + P\big((X = 2) \cap (Y = 3)\big)$

$Y \backslash X$	1	2
-2	-1	0
3	4	5
4	5	6

On cherche toutes les sommes $X + Y$ égales 5.

Si de plus, les évènements X et Y sont indépendants, alors on a :
$$P(X + Y = 5) = P(X = 1)P(Y = 4) + P(X = 2)P(Y = 3)$$

Définition : Soit X et Y deux variables aléatoires. La **loi de probabilité** de la variable aléatoire somme $X + Y$ est donnée par :
$$P(X + Y = k) = \sum_{i+j=k} P((X = i) \cap (Y = j))$$
Si, de plus, les évènements $(X = i)$ et $(Y = j)$ sont indépendants, alors on a :
$$P(X + Y = k) = \sum_{i+j=k} P(X = i)\, P(Y = j)$$

Méthode : Déterminer la loi d'une somme de variables aléatoires

On considère le jeu suivant qui se déroule en deux parties :
- La 1ère partie consiste à lancer une pièce de monnaie. Si on tombe sur « pile », on gagne 1 €, si on tombe sur « face », on gagne 2 €.
- La 2e partie consiste à lancer un dé à 6 faces. Si on tombe sur un chiffre pair, on gagne 1 €, si on tombe sur le « 3 » ou le « 5 », on gagne 2 €.

Si on tombe sur le « 1 », on perd 5 €.
La variable aléatoire X désigne les gains à la 1ère partie, la variable aléatoire Y désigne les gains à la 2e partie.
On considère que les variables aléatoires X et Y sont indépendantes.
Établir la loi de probabilité de la variable aléatoire somme $S = X + Y$ donnant le gain total cumulé à la fin des deux parties.

x_i	1	2
$P(X=x_i)$	0,5	0,5

y_i	1	2	-5
$P(Y=y_i)$	0,5	$\frac{1}{3}$	$\frac{1}{6}$

$X \backslash Y$	1	2	-5
1	2	3	-4
2	3	4	-3

S_i	-4	-3	2	3	4
$P(S=S_i)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{5}{12}$	$\frac{1}{6}$

II. Espérance et variance de combinaisons linéaires de variables aléatoires

1) Linéarité de l'espérance

Propriétés :
1) $E(aX+b) = aE(X) + b$ avec $a \in \mathbb{R}$ et $b \in \mathbb{R}$
2) $E(X+Y) = E(X) + E(Y)$

2) Variance

Propriété :
1) $V(aX+b) = a^2V(X)$ avec $a \in \mathbb{R}$ et $b \in \mathbb{R}$
2) Si X et Y sont deux variables aléatoires indépendantes : $V(X+Y) = V(X) + V(Y)$

Méthode : Simplifier les calculs d'espérance et de variance à l'aide d'une variable aléatoire de transition

Une entreprise fabrique des machines. Soit X la variable aléatoire qui, pour un mois choisi au hasard, associe le nombre de machines vendues pendant cette période.
Un étude statistique permet d'établir la loi de probabilité de X.

x_i	0	1	2	3	4	5	6
$P(X = x_i)$	0.04	0.08	0.12	0.28	0.25	0.17	0.06

1) Calculer l'espérance.

2) La vente d'une machine rapporte 5 000 €. On note Y la variable aléatoire qui, à un mois tiré au hasard, associe le résultat en euros de l'atelier.

Déterminer l'espérance de Y, puis l'interpréter.

3) Le résultat mensuel, en euros, du second atelier de l'entreprise définit une variable aléatoire T d'espérance 20 000.

Quelle est l'espérance du résultat mensuel total de l'entreprise ?

1) $E(X) = 0\times0{,}04 + 1\times0{,}08 + 2\times0{,}12 + 3\times0{,}28 + 4\times0{,}25 + 5\times0{,}17 + 6\times0{,}06$
$= 0{,}08 + 0{,}24 + 0{,}84 + 1 + 0{,}85 + 0{,}36$
$= 3{,}37$

Sur un grand nombre de mois l'entreprise peut espérer vendre 3,37 machines / mois

2)

y_i	0	5000	10 000	15 000	20 000	25 000	30 000
$P(Y=y_i)$	0,04	0,08	0,12	0,28	0,25	0,17	0,06

$Y = 5000X$

$E(Y) = 0\times0{,}04 + 5000\times0{,}08 + 10000\times0{,}12 + 15000\times0{,}28 + 20000\times0{,}25 + 25000\times0{,}17 + 30000\times0{,}06$
$= 16850 = 5000\times3{,}37$

3) $E(Y+T) = E(Y) + E(T) = 16850 + 20000 = 36850$

Capacités 1 et 2 p403 ; 41, 45 ; 48, 54, 55 p412-413

III. Echantillon d'une loi de probabilité

1) Définition

Exemple :
Soit X la variable aléatoire qui, à chaque paquet de chips issues d'une chaine de production, associe sa masse en grammes. On note X_i la variable aléatoire qui, à chaque lot de 3 paquets de chips, associe la masse du i-ème paquet.
On peut considérer alors que la liste (X_1, X_2, X_3) forment un échantillon de taille 3 de variables aléatoires suivant la même loi de X.

Définition : Un **échantillon de taille n d'une loi de probabilité** est une liste de n variables aléatoires indépendantes suivant cette loi.

Définition :
La **variable aléatoire somme** d'un échantillon de taille n de la loi de X est la variable aléatoire définie sur l'ensemble des échantillons de taille n par : $S_n = X_1 + X_2 + \cdots + X_n$.
La **variable aléatoire moyenne** est la variable aléatoire définie par $M_n = \frac{S_n}{n} = \frac{X_1 + \ldots + X_n}{n}$

Exemple :
En reprenant l'exemple précédent,
La variable aléatoire somme S_3 définie par $S_3 = X_1 + X_2 + X_3$ associe à chaque lot sa masse en grammes.
La variable aléatoire somme M_3 définie par $M_3 = \frac{S_3}{3} = \frac{X_1 + X_2 + X_3}{3}$ associe à chaque lot de 3 paquets la masse moyenne d'un paquet.

Propriétés : Soit $(X_1, X_2, \ldots, X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une même loi. On pose : $S_n = X_1 + X_2 + \cdots + X_n$, alors on a :

1) $E(S_n) = n\,E(X)$ et $E(M_n) = E(X)$

2) $V(S_n) = n\,V(X)$ et $V(M_n) = \frac{V(X)}{n}$

3) $\sigma(S_n) = \sqrt{n}\,\sigma(X)$ et $\sigma(M_n) = \frac{\sigma(X)}{\sqrt{n}}$

Méthode : Calculer l'espérance, la variance et l'écart-type d'une variable aléatoire moyenne

On considère la variable aléatoire X qui prend, de façon équiprobable, les valeurs –4, 0, 1, 3 et 6.

M_{50} est la variable aléatoire moyenne d'un échantillon de taille 50 de la loi de X.

Calculer l'espérance, la variance et l'écart type de M_{50}.

x_i	-4	0	1	3	6
$P(X=x_i)$	1/5	1/5	1/5	1/5	1/5

$E(X) = -4 \times \frac{1}{5} + 0 \times \frac{1}{5} + 1 \times \frac{1}{5} + 3 \times \frac{1}{5} + 6 \times \frac{1}{5}$
$= \frac{-4+1+3+6}{5} = \frac{6}{5}$

$V(X) = \left(-4 - \frac{6}{5}\right)^2 \times \frac{1}{5} + \left(-\frac{6}{5}\right)^2 \times \frac{1}{5} \cdots + \left(6 - \frac{6}{5}\right)^2 \times \frac{1}{5}$
$= 10{,}96$

$\sigma(X) = \sqrt{V(X)} = \sqrt{10{,}96}$

$E(M_{50}) = E(X) = \frac{6}{5}$

$V(M_{50}) = 50\,V(X)$
$= 50 \times 10{,}96 = 548$

$\sigma(M_{50}) = \frac{\sigma(X)}{\sqrt{50}} = \frac{\sqrt{10{,}96}}{\sqrt{50}}$
$\simeq 0{,}468$

2) Cas particulier : Loi Binomiale

Propriété : Soit $(X_1, X_2, \dots X_n)$ un échantillon de taille n de la loi de Bernoulli de paramètre p.
La variable aléatoire $S = X_1 + X_2 + \cdots + X_n$ suit la loi binomiale de paramètres n et p.

Propriété : Soit S une variable aléatoire qui suit la loi binomiale de paramètres n et p.
$E(S) = np$ $\qquad V(S) = np(1-p)$ $\qquad \sigma(S) = \sqrt{V(S)}$

Rappel :
Si X suit une loi de Bernoulli de paramètre p, alors $E(X) = p$; $\quad V(X) = p(1-p)$ $\quad$ et $\quad \sigma(X) = \sqrt{p(1-p)}$

Méthode : Calculer l'espérance, la variance et l'écart-type pour une loi binomiale

On lance 5 fois de suite un dé à six faces.
On considère comme succès le fait d'obtenir 5 ou 6.
On considère la variable aléatoire S donnant le nombre de succès.
Calculer $E(S)$, $V(S)$ et $\sigma(S)$.

$S \sim B(\frac{1}{3} ; 5)$

$E(S) = \frac{1}{3} \times 5 = \frac{5}{3}$ $\qquad V(S) = \frac{1}{3} \times 5 \left(1 - \frac{1}{3}\right) = \frac{1}{3} \times 5 \times \frac{2}{3} = \frac{10}{9}$

$\sigma(S) = \sqrt{\frac{10}{9}} = \frac{\sqrt{10}}{\sqrt{9}} = \frac{\sqrt{10}}{3}$

Ex : 65, 66, 68 p415

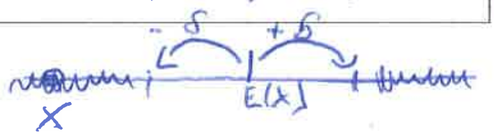
IV- Loi des grands nombres

1) Inégalité de Bienaymé-Tchebychev.

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

$$P(|X - E(X)| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

Plus on s'éloigne de l'espérance plus la proba est petite et est majorée



Méthode : Appliquer l'inégalité de Bienaymé-Tchebychev

Vidéo https://youtu.be/4XMvq1FnYwU

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

1) Appliquer l'inégalité de Bienaymé-Tchebychev avec $\delta = 2\sigma(X)$. Interpréter.

2) Recommencer avec $\delta = 3\sigma(X)$, puis $\delta = 4\sigma(X)$. Que constate-t-on ?

1) $X \sim B(n = 20 ; p = 0{,}1)$

$E(X) = 20 \times 0{,}1 = 2$

$V(X) = 20 \times 0{,}1(1 - 0{,}1) = 2 \times 0{,}9 = 1{,}8$

$\sigma(X) = \sqrt{1{,}8}$

$P(|X - E(X)| \geq \delta) \leq \frac{V(X)}{\delta^2}$

$P(|X - 2| \geq 2\sqrt{1{,}8}) \leq \frac{1{,}8}{(2\sqrt{1{,}8})^2}$

$P(|X - 2| \geq 2\sqrt{1{,}8}) \leq \frac{1}{4}$

la proba que l'écart de X avec $E(X) = 2$ soit $\geq 2\sqrt{1{,}8}$ est majorée par $\frac{1}{4}$

2) $P(|X - 2| \geq 3\sqrt{1{,}8}) \leq \frac{1{,}8}{(3\sqrt{1{,}8})^2} = \frac{1}{9}$ | $P(|X - 2| \geq 4\sqrt{1{,}8}) \leq \frac{1}{16}$

2) Inégalité de concentration

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

$$P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\,\delta^2}$$

Méthode : Appliquer l'inégalité de concentration pour déterminer la taille d'un échantillon

Vidéo https://youtu.be/7Nk9U-zwWOA

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 supérieure à 0,95.

n tel que $P(0{,}03 < M_n < 0{,}37) > 0{,}95$

$E(X) = 0{,}2$

$V(X) = 0{,}2 \times 0{,}8 = 0{,}16$

$\Leftrightarrow P(0{,}03 - 0{,}2 < M_n - 0{,}2 < 0{,}37 - 0{,}2) > 0{,}95$

$\Leftrightarrow P(-0{,}17 < M_n - 0{,}2 < 0{,}17) > 0{,}95$

$\Leftrightarrow P(|M_n - 0{,}2| < 0{,}17) > 0{,}95$

$\Leftrightarrow 1 - P(|M_n - 0{,}2| \geq 0{,}17) > 0{,}95$

$\Leftrightarrow -P(|M_n - 0{,}2| \geq 0{,}17) > -0{,}05$

$\Leftrightarrow P(|M_n - 0{,}2| \geq 0{,}17) < 0{,}05$

Or on sait que $P(|M_n - 0{,}2| \geq 0{,}17) \leq \frac{0{,}16}{0{,}17^2\, n}$

On cherche n tq $\frac{0{,}16}{n\, 0{,}17^2} < 0{,}05 \Leftrightarrow \frac{0{,}16}{0{,}17^2} < 0{,}05\, n \Leftrightarrow \frac{\frac{0{,}16}{0{,}17^2}}{0{,}05} < n$

$n \geq 111$

3) Loi des grands nombres

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

$$\lim_{n \to +\infty} P(|M_n - E(X)| \geq \delta) = 0$$

Remarque : La loi des grands nombres traduit le fait que plus la taille de l'échantillon d'une variable aléatoire X est grande, plus l'écart entre la moyenne de cet échantillon et l'espérance de la variable aléatoire X est faible.

Capacités 5 et 6 p 407 ; 70, 74, 80, 82, 86 p416-417