Complément sur les variables aléatoires, loi des grands nombres

1. Rappels de 1ère : Notion de variable aléatoire

A. Définition

Définition 15.1 Soit Ω l'univers d'une expérience aléatoire (i.e. l'ensemble des issues possibles de cette expérience aléatoire).
Si à chaque éventualité e_i de Ω on associe un réel x_j, on dit que l'on définit une variable aléatoire sur Ω.
Ainsi une variable aléatoire est une fonction définie sur Ω.

Remarques :

- Attention, un même réel peut être associé à plusieurs éventualités.
- Soit X une variable aléatoire associée à une expérience aléatoire. Cette variable permet de définir des événements. Par exemple, s'il s'agit d'un jeu aléatoire dans lequel on gagne des points,
  - $(X = 5)$ est l'événement réalisé lorsque à la fin de la partie, on a gagné 5 points.
  - $(X \leq 10)$ est l'événement réalisé lorsqu'à la fin de la partie on a 10 points ou moins.

Exemple 15.1 On lance un dé à six faces.

- Si le résultat est 6, on gagne 2€
- Si le résultat est pair et inférieur strictement à 6 on ne gagne ni ne perd rien
- Sinon on perd 1€

Soit X la variable aléatoire égale au gain algébrique (c'est-à-dire positif ou négatif) du joueur. On a alors :

Issue	1	2	3	4	5	6
Valeur de X	-1	0	-1	0	-1	2

On remarque que l'événement $(X = 2)$ est le même que l'événement "obtenir un 6". De même, l'événement $(X = 0)$ est le même que l'événement "obtenir un 2 ou un 4". On peut donc définir un événement en utilisant la variable aléatoire X.

B. Loi de probabilité d'une variable aléatoire

Définition 15.2 :
Soit Ω un univers lié à une expérience aléatoire et soit X une variable aléatoire sur Ω prenant pour valeurs $x_1, ..., x_n$. La loi de probabilité de la variable X est la fonction qui à chaque x_i associe la probabilité que la variable prenne cette valeur ; on note cette probabilité $p(X = x_i)$.

Exemple 15.2 La loi de probabilité de la variable aléatoire de l'exemple 15.1 est donnée dans le tableau suivant :

Valeur x_i de X	-1	0	2
$p(X = x_i)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

2. Rappels de 1ère : Paramètres d'une variable aléatoire

A. Définitions : espérance, variance, écart-type

Définition 15.3 Soit une loi de probabilité

x_i	x_1	x_2	...	x_n
$P(X = x_i)$	p_1	p_2	...	p_n

- On appelle espérance le réel $E(X)$ défini par :

$$E(X) = \sum_{i=1}^{n} p_i x_i$$

- On appelle variance le réel positif $V(X)$ défini par :

$$V(X) = \sum_{i=1}^{n} p_i \Big(x_i - E(X)\Big)^2$$

- On appelle écart-type le réel positif $\sigma(X)$ défini par :

$$\sigma(X) = \sqrt{V(X)}$$

Interprétation et "trucs" mnémotechniques :
L'espérance mathématique d'une variable aléatoire réelle est l'équivalent en probabilité de ce qu'est la moyenne d'une série statistique en statistiques.
La variance et l'écart-type sont des mesures de dispersion.
La variance peut être vue comme "la moyenne des carrés des écarts à la moyenne" (rigoureusement : l'espérance des carrés des écarts à l'espérance).
La deuxième expression de la variance proposée dans la propriété peut être vue comme la "moyenne des carrés moins le carré de la moyenne".

Exemple 15.3 Reprise de l'exemple 15.1 .

$E(X) = (-1) \times \frac{1}{2} + 0 \times \frac{1}{3} + 2 \times \frac{1}{6} = -\frac{1}{6}$ $\quad V(X) = \frac{41}{36}$ $\quad \sigma(X) = \sqrt{\frac{41}{36}} = \frac{\sqrt{41}}{36}$

Remarques :

- Dans un jeu où X est associé aux gains possibles, $E(X)$ représente le gain moyen que l'on peut espérer. Le jeu est dit équitable si $E(X) = 0$.
- Si on réalisait une expérience aléatoire un très grand nombre de fois, les résultats statistiques (fréquences, moyenne, écart-type) se rapprocheraient des résultats trouvés via le modèle théorique probabiliste (probabilités, espérance, écart-type).

B. Propriétés

Soit X une variable aléatoire. Cette variable X prenant des valeurs réelles, on peut effectuer des opérations sur elle (la multiplier par un réel, lui appliquer une fonction, . . .). En reprenant les données de l'exemple précédent, on peut choisir de définir une nouvelle variable aléatoire Y qui est le résultat d'une multiplication des gains algébriques par 2. On écrit alors $Y = 2X$. Et on aura :

Valeur y_i de Y	-2	0	4
$p(Y = y_i)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

On a alors $E(Y) = (-2) \times \frac{1}{2} + 0 \times \frac{1}{3} + 4 \times \frac{1}{6} = -\frac{1}{3}$.

Propriété 15.1 Soit X une variable aléatoire et soient a et b deux réels. On a :
$E(aX + b) = aE(X) + b$, $\quad V(aX + b) = a^2V(X)$ et $\quad \sigma(aX + b) = |a|\sigma(X)$

Démonstration

$$\begin{aligned}
E(aX+b) &= (ax_1+b)\times p_1 + (ax_2+b)\times p_2 + ... + (ax_n+b)\times p_n \\
&= (ax_1p_1 + ax_2p_2 + ... + ax_np_n) + (bp_1 + bp_2 + ... + bp_n) \\
&= a(x_1p_1 + x_2p_2 + ... + x_np_n) + b(p_1 + p_2 + ... + p_n) \\
&= aE(X) + b\times 1 \\
&= aE(X) + b
\end{aligned}$$

$$\begin{aligned}
V(aX+b) &= \sum_{i=1}^{n} p_i(ax_i + b - E(aX+b))^2 \\
&= \sum_{i=1}^{n} p_i(ax_i + b - aE(X) - b)^2 \\
&= \sum_{i=1}^{n} p_i(ax_i - aE(X))^2 \\
&= \sum_{i=1}^{n} p_i\left(a(x_i - E(X))\right)^2 \\
&= \sum_{i=1}^{n} p_i \times a^2 \times (x_i - E(X))^2 \\
&= a^2 \times \sum_{i=1}^{n} p_i \times (x_i - E(X))^2 \\
&= a^2V(X)
\end{aligned}$$

Propriété 15.2 Soient X et Y deux variables aléatoires. On a :
$E(X+Y) = E(X) + E(Y)$

Propriété 15.3 (Application) Soit X une variable aléatoire d'espérance $E(X)$ et de variance $V(X)$.
Alors on a :

$$V(X) = E(X^2) - \Big(E(X)\Big)^2$$

Démonstration

$$\begin{aligned}
V(X) &= \sum_{i=1}^{n} p_i\Big(x_i - E(X)\Big)^2 \\
&= \sum_{i=1}^{n} p_i x_i^2 - 2\sum_{i=1}^{n} p_i x_i E(X) + \sum_{i=1}^{n} p_i\Big(E(X)\Big)^2 \\
&= E(X^2) - 2\sum_{i=1}^{n} \mathbf{p_i} x_i \mathbf{E}(\mathbf{X}) + \sum_{i=1}^{n} \mathbf{p_i} E(X) \times \mathbf{E}(\mathbf{X}) \\
&= E(X^2) - \sum_{i=1}^{n} \Big(\mathbf{p_i}\mathbf{E}(\mathbf{X})\,(2x_i - E(X))\Big) \\
&= E(X^2) - E(X)\left(\sum_{i=1}^{n} 2p_i x_i - \sum_{i=1}^{n} p_i E(X)\right) \\
&= E(X^2) - E(X)\left(2E(X) - E(X)\sum_{i=1}^{n} p_i\right) \\
&= E(X^2) - E(X)\,(2E(X) - E(X) \times 1) \\
&= E(X^2) - E(X) \times E(X) \\
&= E(X^2) - E(X)^2
\end{aligned}$$

Exemple 15.4 Toujours avec l'exemple 15.1, en posant $Z = X^2$, on a :

Valeur z_i de Z	1	0	4
$p(Z = z_i)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

Donc $E(Z) = 1 \times \frac{1}{2} + 0 \times \frac{1}{3} + 4 \times \frac{1}{6} = \frac{7}{6}$.

Ainsi on a :

$$V(X) = E(Z) - \Big(E(X)\Big)^2 = \frac{7}{6} - \left(-\frac{1}{6}\right)^2 = \frac{7}{6} - \frac{1}{36} = \frac{41}{36}$$

On retrouve bien le résultat de l'exemple 15.3.

C. Variables aléatoires indépendantes

Deux variables aléatoires sont dites ***indépendantes*** lorsqu'elles sont associées à des épreuves indépendantes.

Propriété 15.4 Caractérisation Deux variables aléatoires X et Y sont ***indépendantes si et seulement si*** quelles que soient les valeurs x_i et y_j,

$$P(X = x_i \text{ et } Y = y_j) \ = \ P(X = x_i) \times P(Y = y_j)$$

Propriété 15.5 Soient X et Y deux variables aléatoires ***indépendantes*** .
Alors $V(X + Y) = V(X) + V(Y)$.

3. Échantillon d'une loi de probabilité

A. Échantillon de taille n d'une loi de probabilité

Définition 15.4 soit X une v.a. définie sur Ω. ***Un échantillon de taille n*** de la loi de X est une liste $(X_1, X_2, ..., X_n)$ de v.a. indépendantes et identiques suivant cette loi.

Exemple 15.5 Dans le cas d'un lancer de dés, un échantillon de taille 4 correspond à la liste des 4 résultats de 4 lancers du dé.

Définition 15.5 • La ***variable aléatoire somme*** d'un échantillon de taille n de la loi X est la v.a. définie sur l'ensemble des échantillons de taille n par

$$S_n = X_1 + X_2 + ... + X_n$$

• La ***variable aléatoire moyenne*** est le v.a. définie par $M_n = \frac{1}{n}S_n$

Exemple 15.6 Dans le cas d'un lancer de dés, la v.a. somme est la somme des 4 résultats, et la v.a. moyenne est la moyenne des 4 résultats.

B. Espérance, variance et écart-type de S_n et M_n

Propriété 15.6 Soit S_n la v.a. somme d'un échantillon de taille n de la v.a. X, alors :
$E(S_n) = nE(X)$, $V(S_n) = nV(X)$ et $\sigma(S_n) = \sqrt{n}\sigma(X)$

Démonstration $\forall k \in [\![1;n]\!]$, $E(X_k) = E(X)$ et $V(X_k) = V(X)$.
Par suite $E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$, d'où $E(S_n) = nE(X)$.
De plus, $X_1, X_2, ...X_n$ sont des v.a. indépendantes, donc :
$V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n)$
Finalement, $V(S_n) = nV(X)$, et $\sigma(S_n) = \sqrt{nV(X)} = \sqrt{n}\sqrt{V(X)} = \sqrt{n}\sigma(X)$.

Propriété 15.7 Soit M_n la v.a. moyenne d'un échantillon de taille n de la v.a. X ; alors :
$E(M_n) = E(X)$, $V(M_n) = \frac{1}{n}V(X)$, et $\sigma(S_n) = \frac{\sigma(X)}{\sqrt{n}}$

Remarque 15.1 On en déduit l'espérance et la variance de la loi binomiale, vues au chapitre 9.

4. Loi des grands nombres

A. Espérance, variance et écart-type de S_n et M_n

L'écart-type σ représente (mais n'est pas strictement égal à) l'écart entre les valeurs prises par la v.a. et l'espérance de celle-ci : la variance est la moyenne du carré des écarts à la moyenne, et l'écart type est la racine carrée de la variance.
L'inégalité de Bienaymé-Tchebychev précise cette idée.

Théorème 15.1 Inégalité de Bienaymé-Tchebychev.
Soit X une v.a. d'espérance μ et de variance V.
$\forall \delta \in \mathbb{R}_+^*$, $P\left(|X - \mu| \geqslant \delta\right) \leqslant \frac{V(X)}{\delta^2}$

Démonstration Soit $\Omega' = \{x_1; x_2; ...x_n\}$ l'ensemble des valeurs prises par la v.a. X, et soit l'événement $A = \{|X - \mu| > \delta\}$, dont on note les éléments $a_1, a_2, ...a_k$.
Ces éléments $a_1, a_2, ...a_k$ appartiennent à A, et $\forall i \in [\![1;k]\!]$, $|a_i - \mu| > \delta$.
On sait que $V(X) = (x_1 - \mu)^2 P(X = x_1) + ... + (x_n - \mu)^2 P(X = x_n)$.
En séparant dans cette somme les termes relatifs à des éléments de A des autres termes, il vient :
$$V(X) = \underbrace{\sum_{x_p \in A} (x_p - \mu)^2 P(X = x_p)}_{S} + \underbrace{\sum_{x_q \notin A} (x_n - \mu)^2 P(X = x_n)}_{S'}$$
Où tous les termes composant la somme S' sont positifs, donc : $V(X) \geqslant \sum\limits_{x_p \in A} (x_p - \mu)^2 P(X = x_p)$,
c'est-à-dire, avec les notations précédentes pour les éléments de A :
$V(X) \geqslant (a_1 - \mu)^2 P(X = a_1) + ... + (a_k - \mu)^2 P(X = a_k)$
Par définition de A, chacun des nombres $|a_i - \mu|$ est supérieur à δ, donc chacun des nombres $(a_i - \mu)^2$ est supérieur à δ^2. Il vient :
$(a_1 - \mu)^2 P(X = a_1) + ... + (a_k - \mu)^2 P(X = a_k) \geqslant \delta^2 P(X = a_1) + ... + \delta^2 P(X = a_k)$
D'où : $V(X) \geqslant \delta^2 P(X = a_1) + ... + \delta^2 P(X = a_k)$. En factorisant δ^2 :
$V(X) \geqslant \delta^2 \left(P(X = a_1) + ... + P(X = a_k)\right)$, avec $P(X = a_1) + ... + P(X = a_k) = P(X \in A)$.
Ainsi, $V(X) \geqslant \delta^2 P(X \in A)$, i.e. $P\left(X \in A\right) \leqslant \frac{V(X)}{\delta^2}$, ce qui peut encore s'écrire $P\left(|X - \mu| \geqslant \delta\right) \leqslant \frac{V(X)}{\delta^2}$

Propriété 15.8 Inégalité de concentration.
Soit M_n la v.a. moyenne d'un échantillon de taille n d'une v.a. d'espérance μ et de variance V.
Alors, pour tout réel $\delta > 0$, on a : $P(|M_n - \mu| \geqslant \delta) \leqslant \frac{V}{n\delta^2}$.

Démonstration On applique l'inégalité de Bienaymé-Tchebychev à M_n, il vient :

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

Or $E(M_n) = \mu$, et comme $V(M_n) = \frac{V}{n}$, il vient $P(|M_n - \mu| \geqslant \delta) \leqslant \frac{V}{n\delta^2}$.

B. Loi faible des grands nombres

Théorème 15.2 Soit M_n la v.a. moyenne d'un échantillon de taille n d'une v.a. d'espérance μ et de variance V.
Alors, pour tout réel $\delta > 0$, on a : $\lim\limits_{n\to+\infty} P(|M_n - \mu| \geqslant \delta) = 0$.

Démonstration Soit δ un réel strictement positif. D'après l'inégalité de concentration, pour tout entier naturel n non nul, $P(|M_n - \mu| \geqslant \delta) \leqslant \frac{V}{n\delta^2}$.
D'où $0 \leqslant P(|M_n - \mu| \geqslant \delta) \leqslant \frac{V}{n\delta^2}$, et comme $\lim\limits_{n\to+\infty} \left(\frac{V}{n\delta^2}\right) = 0$, en appliquant le théorème des gendarmes il vient $\lim\limits_{n\to+\infty} P(|M_n - \mu| \geqslant \delta) = 0$.

La probabilité que la moyenne s'écarte de l'espérance devient très petite pour les grands échantillons.
On pourra donc "modéliser" certaines probabilités en se basant sur les fréquences obtenues pour de très grands échantillons.