Somme de variables aléatoires et concentration

I. Variables aléatoires et opérations

Soit X et Y deux variables aléatoires définies sur l'univers Ω et w les éléments de cet univers. On peut alors définir une variable aléatoire Z sur Ω telle que, pour tout élément $w \in \Omega$, $\mathrm{Z}(w)=\mathrm{X}(w)+\mathrm{Y}(w)$.
La variable aléatoire Z est appelée **somme de variable aléatoires** X et Y. On note $\mathrm{Z}=\mathrm{X}+\mathrm{Y}$.

Soit X une variable aléatoire définie sur l'univers Ω, a un nombre réel et w les éléments de l'univers Ω. On peut alors définir une variable aléatoire Y telle que, pour tout élément $\omega \in \Omega$, $\mathrm{Y}(w)=a\,\mathrm{X}(w)$. On note $\mathrm{Y}=a\,\mathrm{X}$ cette variable aléatoire.

II. Espérance et variance d'une somme de variables aléatoires

II.1 Espérance

Soit X et Y deux variables aléatoires définies sur le même univers Ω.
Alors on a $\mathrm{E}(\mathrm{X}+\mathrm{Y})=\mathrm{E}(\mathrm{X})+\mathrm{E}(\mathrm{Y})$.

Remarques
• On peut déterminer l'espérance d'une addition de plus de deux variables aléatoires. Si Z est aussi une variable aléatoire, alors $\mathrm{E}(\mathrm{X}+\mathrm{Y}+\mathrm{Z})=\mathrm{E}(\mathrm{X})+\mathrm{E}(\mathrm{Y})+\mathrm{E}(\mathrm{Z})$.
• Si Y est constante égale à b, où b est un réel, alors $\mathrm{E}(\mathrm{X}+b)=\mathrm{E}(\mathrm{X})+b$ car $\mathrm{E}(b)=b$.

Soit X et Y deux variables aléatoires définies sur un même univers Ω et a un nombre réel.
Alors on a $\mathrm{E}(a\,\mathrm{X})=a\,\mathrm{E}(\mathrm{X})$.

Remarques
• Plus généralement, on a $\mathrm{E}(a\,\mathrm{X}+b\,\mathrm{Y})=a\,\mathrm{E}(\mathrm{X})+b\,\mathrm{E}(\mathrm{Y})$.
• On appelle les propriétés 12 et 13 la **linéarité de l'espérance**.

II.2 Variance

Soit X_1, X_2, …, X_n, n variables aléatoires à valeurs respectivement dans E_1, E_2, …, E_n.
On dit que X_1, X_2, …, X_n sont indépendantes lorsque, pour tous $x_1 \in \mathrm{E}_1$, $x_2 \in \mathrm{E}_2$, …, $x_n \in \mathrm{E}_n$, on a :

$$\mathrm{P}(\mathrm{X}_1=x_1 \cap \mathrm{X}_2=x_2 \cap \ldots \cap \mathrm{X}_n=x_n)=\mathrm{P}(\mathrm{X}_1=x_1)\times\mathrm{P}(\mathrm{X}_2=x_2)\times\ldots\times\mathrm{P}(\mathrm{X}_n=x_n)$$

Soit X et Y deux variables aléatoires **indépendantes** définies sur l'univers Ω.
Alors on a $\mathrm{V}(\mathrm{X}+\mathrm{Y})=\mathrm{V}(\mathrm{X})+\mathrm{V}(\mathrm{Y})$.

Soit X une variable aléatoire définie sur l'univers Ω dont on note $V(X)$ sa variance.
Alors on a $V(aX)=a^2V(X)$ et $\sigma(aX)=|a|\sigma(X)$.

Remarques
• De manière intuitive, deux variables aléatoires sont indépendantes si les résultats de l'une n'ont pas influence sur les résultats de l'autre.
• Si X_1, X_2, …, X_n sont des variables aléatoires indépendantes, alors on a
$V(X_1+X_2+\ldots+X_n)=V(X_1)+V(X_2)+\ldots+V(X_n)$.

III. Applications

III.1 Loi binomiale

Deux variables aléatoires sont **identiquement distribuées** lorsqu'elles ont la même loi de probabilité.

Soit X_1, X_2, … ; X_n, n variables aléatoires indépendantes suivant toutes une même loi de Bernoulli de paramètre p.
Alors la variable aléatoire $X_1+X_2+\ldots+X_n$ suit la loi binomiale de paramètres n et p.

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p.
Alors on a $X=X_1+X_2+\ldots X_n$ où les variables X_1, X_2, …, X_n sont des variables aléatoires de Bernoulli indépendantes et identiquement distribuées.

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p.
• L'espérance de X est $E(X)=np$.
• La variance de X est $V(X)=np(1-p)$.
• L'écart-type de X est $\sigma(X)=\sqrt{np(1-p)}$.

III.2 Échantillonnage de n variables aléatoires identiques et indépendantes

Une liste $(X_1;X_2;\ldots;X_n)$ de variables aléatoires indépendantes définies sur l'univers Ω suivant toutes la même loi est appelée **échantillon de taille n** associée à cette loi (ou associée à une variable aléatoire X suivant cette loi).

On considère n expériences aléatoires identiques et indépendantes. On note X_1, X_2, …, X_n les variables aléatoires associées à ces expériences, toutes de même loi.
On appelle respectivement S_n et M_n les **variables aléatoires somme et moyenne** d'un échantillon de taille n des X_i. On a $S_n=\sum_{i=1}^{n}X_i$ et $M_n=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{S_n}{n}$.

Soit $(X_1 ; X_2 ; \ldots ; X_n)$ un échantillon de taille n d'une variable aléatoire X.
- Les espérances de S_n et M_n sont $E(S_n) = n\,E(X_i)$ et $E(M_n) = E(X_i)$.
- Les variances de S_n et M_n sont $V(S_n) = n\,V(X_i)$ et $V(M_n) = \dfrac{V(X_i)}{n}$.
- Les écarts-types de S_n et M_n sont $\sigma(S_n) = \sqrt{n}\,\sigma(X_i)$ et $\sigma(M_n) = \dfrac{\sigma(X_i)}{\sqrt{n}}$.

IV. Inégalités de concentration

IV.1 Inégalité de Markov

Une variable aléatoire est dite positive ou nulle dans un univers Ω lorsque toutes les valeurs prises par celle-ci sont des réels positifs ou nuls.

Inégalité de Markov
Soit X une variable aléatoire réelle à valeurs positives ou nulles d'espérance $E(X)$.
Pour tout réel δ strictement positif, on a $P(X \geqslant \delta) \leqslant \dfrac{E(X)}{\delta}$.

IV.2 Inégalité de Bienaymé-Tchebychev

Inégalité de Bienaymé-Tchebychev
Soit X une variable aléatoire réelle d'espérance $E(X) = \mu$ et de variance $V(X)$.
Pour tout réel δ strictement positif, on a $P(|X - \mu| \geqslant \delta) \leqslant \dfrac{V(X)}{\delta^2}$.

IV.3 Inégalité de concentration

Inégalité de concentration
Soit X_1, X_2, …, X_n un échantillon de variables aléatoires d'espérance μ et de variance V et M_n la variable aléatoire moyenne de cet échantillon.
Pour tout réel δ strictement positif, on a $P(|M_n - \mu| \geqslant \delta) \leqslant \dfrac{V}{n\delta^2}$.

V. Loi des grands nombres

Soit X_1, X_2, …, X_n un échantillon de variables aléatoires d'espérance μ et de variance $V(X)$ et M_n la variable aléatoire moyenne de cet échantillon.
Pour tout réel δ strictement positif, on a $\lim\limits_{n \to +\infty} P(|M_n - \mu| \geqslant \delta) = 0$.
On dit que M_n **converge en probabilité** vers μ lorsque n tend vers $+\infty$.