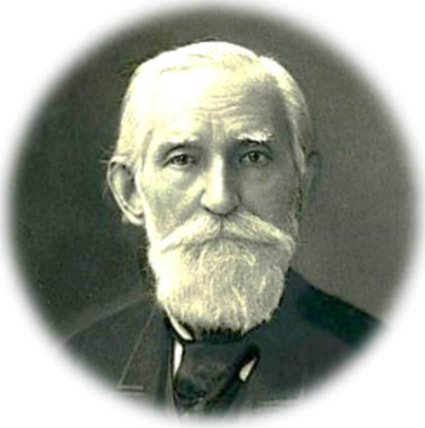
Irénée-Jules Bienaymé (1796 - 1878)
Pafnouti Tchebychev (1821 - 1894)



Le premier est un probabiliste et statisticien français. Il contribue à la théorie des probabilités, au développement de la statistique et à leurs applications aux calculs financiers, à la démographie et aux statistiques sociales.
Le second est un mathématicien russe connu pour ses travaux dans les domaines des probabilités, des statistiques, et de la théorie des nombres.

I. Somme de variables aléatoires indépendantes

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 :

$$P(X+Y=k)=\sum_{i+j=k} P(X=i)\times P(Y=j)$$

On dit dans ce cas que les **variables aléatoires *X* et *Y* sont indépendantes**.

Exemple 1 :

On lance deux dés, l'un octaédrique dont les faces sont numérotées de 1 à 8 et l'autre tétraédrique dont les faces sont numérotées de 1 à 4.

Soit X et Y les variables aléatoires associées respectivement aux deux dés.

1. Déterminer la loi de $X + Y$.
2. Comparer $E(X + Y)$ avec $E(X)$ et $E(Y)$
3. Comparer $V(X + Y)$ avec $V(X)$ et $V(Y)$

	$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{4}$	$\boxed{5}$	$\boxed{6}$	$\boxed{7}$	$\boxed{8}$
$\boxed{1}$	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
$\boxed{2}$	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\boxed{3}$	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
$\boxed{4}$	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**

La loi de $X + Y$ est donc :

$\boldsymbol{X+Y=k}$	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
$\boldsymbol{P(X+Y=k)}$	$\boldsymbol{\frac{1}{32}}$	$\boldsymbol{\frac{2}{32}}$	$\boldsymbol{\frac{3}{32}}$	$\boldsymbol{\frac{4}{32}}$	$\boldsymbol{\frac{4}{32}}$	$\boldsymbol{\frac{4}{32}}$	$\boldsymbol{\frac{4}{32}}$	$\boldsymbol{\frac{4}{32}}$	$\boldsymbol{\frac{3}{32}}$	$\boldsymbol{\frac{2}{36}}$	$\boldsymbol{\frac{1}{32}}$

$$E(X+Y) = \frac{2 + 2 \times 3 + 3 \times 4 + 4 \times (5 + 6 + 7 + 8 + 9) + 3 \times 10 + 2 \times 11 + 12}{32} = \frac{224}{32} = 7$$

$$V(X+Y) = \frac{2^2 + 2 \times 3^2 + 3 \times 4^2 + 4 \times (5^2 + 6^2 + 7^2 + 8^2 + 9^2) + 3 \times 10^2 + 2 \times 11^2 + 12^2}{32} - 7^2$$

$$V(X+Y) = \frac{1776}{32} - 7^2 = 55{,}5 - 49 = 6{,}5$$

X suit une loi uniforme sur $\{1; 2; 3; 4; 5; 6; 7; 8\}$
i.e. chaque valeur est équiprobable donc de probabilité 1/8 ici

$$E(X) = \frac{1+2+3+4+5+6+7+8}{8} = \frac{36}{8} = 4{,}5$$

$$V(X) = \frac{1^2+2^2+3^2+4^2+5^2+6^2+7^2+8^2}{8} - 4{,}5^2 = \frac{204}{8} - 20{,}25 = 5{,}25$$

Y suit une loi uniforme sur $\{1; 2; 3; 4\}$

$$E(Y) = \frac{1+2+3+4}{4} = \frac{10}{4} = 2{,}5$$

$$V(Y) = \frac{1^2+2^2+3^2+4^2}{4} - 2{,}5^2 = \frac{30}{4} - 6{,}25 = 1{,}25$$

On remarque donc que $\boldsymbol{E(X+Y) = E(X) + E(Y)}$

et que $\boldsymbol{V(X+Y) = V(X) + V(Y)}$ *(attention on ne peut pas en déduire que les VA, X et Y, sont indépendantes)*

Propriétés :

Soient $a \in \mathbb{R}$ et $b \in \mathbb{R}$,

$E(aX + b) = aE(X) + b$

$E(X + Y) = E(X) + E(Y)$

$V(aX + b) = a^2V(X)$

Propriétés :

Si X et Y sont deux variables aléatoires indépendantes alors $V(X + Y) = V(X) + V(Y)$

Définition :

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

Propriétés :

Soit $(X_1, X_2, \ldots, X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une même loi.

$$E(X_1 + X_2 + \cdots + X_n) = nE(X_1)$$
$$V(X_1 + X_2 + \cdots + X_n) = nV(X_1)$$
$$E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = E(X_1)$$
$$V\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n}V(X_1)$$

Interprétation :

La variance de la variable aléatoire moyenne est plus faible que la variance de la variable initiale.

En particulier, plus la taille de l'échantillon n augmente, plus la dispersion de la variable aléatoire moyenne diminue, c'est-à-dire que les valeurs de la variable aléatoire moyenne sont de plus en plus proches de l'espérance.

Exemple 2 :
On tire trois cartes dans un jeu de 32 cartes, **avec remise** entre chaque carte tirée. On suppose chaque tirage indépendant.
On gagne 4 euros pour chaque as tiré, 1 euro pour chaque figure obtenue mais on perd 2 euros pour toutes les autres cartes.

On note G la variable aléatoire donnant le gain algébrique total.

Décomposer G sous la forme d'une combinaison linéaire de trois variables aléatoires que l'on définira puis calculer son espérance.

Soit X_1 le nombre d'as obtenus, X_2 le nombre de figures et X_3 le nombre de cartes restantes.

La probabilité de tirer un as est $\frac{4}{32} = \frac{1}{8}$.
La probabilité de tirer une figure est $\frac{4+4+4}{32} = \frac{3}{8}$.
La probabilité de tirer une carte restante est $\frac{32-16}{32} = \frac{1}{2}$.

On répète 3 fois, de façon identique et indépendante, l'alternative où la probabilité de tirer un as est $\frac{1}{8}$, donc X_1 le nombre d'as suit une loi binomiale de paramètres $n = 3$ et $p = \frac{1}{8}$.

De la même façon :

$$X_2 \hookrightarrow \mathcal{B}\left(n = 3; p = \frac{3}{8}\right) \quad \text{et} \quad X_3 \hookrightarrow \mathcal{B}\left(n = 3; p = \frac{1}{2}\right)$$

J'en déduis que $E(X_1) = 3 \times \frac{1}{8} = \frac{3}{8}$

$$E(X_2) = 3 \times \frac{3}{8} = \frac{9}{8} \quad \text{et} \quad E(X_3) = 3 \times \frac{1}{2} = \frac{3}{2}$$

De plus $G = 4 \times X_1 + 1 \times X_2 - 2 \times X_3$ donc

$$E(G) = E(4X_1 + X_2 - 2X_3)$$

$$= 4E(X_1) + E(X_2) - 2E(X_3)$$

$$= 4 \times \frac{3}{8} + \frac{9}{8} - 2 \times \frac{3}{2}$$

$$E(G) = -0{,}375\ €$$

II. Loi des grands nombres

Une inégalité de concentration fournit des bornes sur la probabilité qu'une variable aléatoire dévie d'une certaine valeur.

Théorème : Inégalité de Bienaymé-Tchebychev

Pour une variable aléatoire X d'espérance μ et de variance V, et quel que soit le réel strictement positif t :

$$P(|X-\mu| \geq t) \leq \frac{V}{t^2}.$$

Cette inégalité indique que la probabilité qu'une variable aléatoire dévie d'une valeur de plus en plus grande de sa moyenne est de plus en plus faible.

Exemple 3 :

On lance 300 fois une pièce équilibrée. On appelle X la variable aléatoire qui donne le nombre de Face obtenus.

a) Déterminer la loi de X, on précisera ses paramètres.

b) Déterminer son espérance et sa variance.

c) Donner une minoration de la probabilité $P(140 < X < 160)$

a) On répète 300 fois, de façon identique et indépendante, l'alternative où la probabilité d'obtenir Face est 0,5, donc X le nombre de Face obtenus suit une loi binomiale de paramètres $n = 300$ et $p = \frac{1}{2}$.

b) $E(X) = np = 300 \times \frac{1}{2} = 150$

$$V(X) = np(1-p) = 300 \times \frac{1}{2} \times \frac{1}{2} = 75$$

c) $$\begin{aligned} P(140 < X < 160) &= P(-10 < X - 150 < 10) \\ &= P(|X - 150| < 10) \\ &= 1 - P(|X - 150| \geq 10) \end{aligned}$$

$$\forall t > 0, \qquad P(|X-\mu| \geq t) \leq \frac{V}{t^2}$$

$$P(|X-150| \geq 10) \leq \frac{75}{10^2}$$

$$P(|X-150| \geq 10) \leq 0{,}75$$

$$-P(|X-150| \geq 10) \geq -0{,}75$$

$$1-P(|X-150| \geq 10) \geq 1-0{,}75$$
$$P(140 < X < 160) \geq 0{,}25$$

Théorème : Loi des grands nombres

Si M_n est la variable aléatoire moyenne d'un échantillon de taille n d'une variable aléatoire d'espérance μ et de variance V, alors pour tout $t > 0$,

$$P(|M_n - \mu| \geq t) \leq \frac{V}{n\, t^2}$$

Pour tout $t > 0$,

$$\lim_{n \to +\infty} P(|M_n - \mu| \geq t) = 0$$

La loi des grands nombres établit 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.

Exemple 4 :

On considère une variable aléatoire X qui suit la loi de Bernoulli de paramètre 0,5 ainsi qu'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 pour que la probabilité que la moyenne M_n appartienne à l'intervalle $]0{,}45; 0{,}55[$ soit supérieure à 0,95.

La loi de X est :

$$P(X=0)=\frac{1}{2}\ \text{ et }\ P(X=1)=\frac{1}{2}$$

$$E(X)=0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2}=\mu$$

$$E(X^2)=0^2\times\frac{1}{2}+1^2\times\frac{1}{2}=\frac{1}{2}$$

$$V(X)=E(X^2)-E(X)^2=\frac{1}{2}-\left(\frac{1}{2}\right)^2=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}=V$$

$$P(M_n\in\,]0{,}45;0{,}55[)=P(0{,}45<M_n<0{,}55)$$

$$= P\left(0{,}45 - \frac{1}{2} < M_n - \mu < 0{,}55 - \frac{1}{2}\right)$$

$$= P(-0{,}05 < M_n - \mu < 0{,}05)$$

$$= P(|M_n - \mu| < 0{,}05)$$

$$= 1 - P(|M_n - \mu| \geq 0{,}05)$$

On souhaite que

$$P(M_n \in]0{,}45; 0{,}55[) \geq 0{,}95$$

$$\Leftrightarrow 1 - P(|M_n - \mu| \geq 0{,}05) \geq 0{,}95$$

$$\Leftrightarrow -P(|M_n - \mu| \geq 0{,}05) \geq -0{,}05$$

$$\Leftrightarrow P(|M_n - \mu| \geq 0{,}05) \leq 0{,}05$$

Or d'après la loi des grands nombres, pour tout $t > 0$,

$$P(|M_n - \mu| \geq t) \leq \frac{V}{n\,t^2}$$

Il suffit donc de choisir n tel que

$$\frac{V}{n\,t^2} \leq 0{,}05 \quad \text{où } t = 0{,}05$$

$$\frac{1/4}{n\,0{,}05^2} \leq 0{,}05$$

$$\Leftrightarrow \frac{0{,}25}{0{,}05 \times \ 0{,}05^2} \leq n$$

$$\Leftrightarrow n \geq 2\ 000$$