$$E(X) = \frac{1}{8}(0+1+1+2+1+2+2+3) = \frac{3}{2}$$

Le calcul peut paraître plus long mais dans certaines situations, il peut s'avérer plus pratique (voir par exemple la démonstration de la linéarité de l'espérance en 2.4.)

<u>Exercice théorique</u> : démontrer que l'espérance $E(X)$ minimise la fonction f définie sur $\mathbb{R}$ par :

$$f(x) = \sum_{i=1}^{n} p_i (x_i - x)^2$$

mais pas la fonction g définie par :

$$g(x) = \sum_{i=1}^{n} p_i |x_i - x|$$

f est la fonction "moyenne des carrés des écarts" tandis que g est la fonction "moyenne des écarts".

La fonction f est dérivable comme somme de fonctions dérivables et on a pour tout $x \in \mathbb{R}$:

$$f'(x) = -2\sum_{i=1}^{n} p_i (x_i - x) = -2\sum_{i=1}^{n} p_i x_i - 2x\sum_{i=1}^{n} p_i = -2(E(X) - x)$$

On en déduit : $$f'(x) \geqslant 0 \Leftrightarrow x \geqslant E(X)$$

Donc f admet un minimum en $E(X)$ (et ce minimum est $f(E(X)) = V(X)$ ...)

L'espérance est donc la quantité qui minimise la moyenne des carrés des écarts. Par contre, elle ne minimise pas la moyenne des écarts. En effet, considérons la variable aléatoire X définie par la loi suivante :

x_i	0	1000
p_i	0,9	0,1

On a : $$E(X) = p_1 x_1 + p_2 x_2 = 100$$

$$g(E(X)) = p_1|x_1 - 100| + p_2|x_2 - 100| = 90 + 90 = 180$$

Or : $$g(0) = E(X) = 100$$

Donc : $$g(0) < g(E(X))$$

Conclusion : $E(X)$ ne minimise pas la fonction g

Quelle est donc la quantité qui minimise la fonction g ? Etudions ça de près.

Quitte à réindexer les indices des x_i, on peut supposer sans perte de généralité que :

$$x_1 < x_2 < \ldots < x_n$$

Donnons-nous $k \in [\![1, n-1]\!]$ et $x \in [x_k, x_{k+1}]$. En coupant la somme en deux, il vient :

$$g(x) = \sum_{i=1}^{n} p_i |x_i - x| = \sum_{i=1}^{k} p_i (x - x_i) + \sum_{i=k+1}^{n} p_i (x_i - x) = x\sum_{i=1}^{k} p_i - \sum_{i=1}^{k} p_i x_i + \sum_{i=k+1}^{n} p_i x_i - x\sum_{i=k+1}^{n} p_i$$

Donc g est affine sur $[x_k, x_{k+1}]$: $$g(x) = a_k x + b_k$$

avec : $$a_k = \sum_{i=1}^{k} p_i - \sum_{i=k+1}^{n} p_i = 2\sum_{i=1}^{k} p_i - 1 \quad \text{et} \quad b_k = \sum_{i=k+1}^{n} p_i x_i - \sum_{i=1}^{k} p_i x_i$$

Elle est donc croissante sur $[x_k, x_{k+1}]$ si et seulement si $a_k \geqslant 0$, c'est-à-dire si et seulement si $\sum_{i=1}^{k} p_i \geqslant \frac{1}{2}$.

Donc g est une fonction affine par morceaux, décroissante sur les intervalles $[x_k, x_{k+1}]$ tels que $\sum_{i=1}^{k} p_i \leqslant \frac{1}{2}$ et croissante sur les intervalles $[x_k, x_{k+1}]$ tels que $\sum_{i=1}^{k} p_i \geqslant \frac{1}{2}$.

Elle admet donc un minimum en la valeur médiane de la série des x_i ($1 \leqslant i \leqslant n$).

2.4. Théorème *Linéarité de l'espérance*

Soient X et Y deux variables aléatoires définies sur le même univers Ω de cardinal fini.

Soit P une probabilité sur Ω.

On a : $$E(X + Y) = E(X) + E(Y)$$

Et en particulier, pour tout réel b : $$E(X + b) = E(X) + b$$

$$E(kX) = kE(X) \text{ pour tout réel } k$$

Démonstration :

$$E(X + Y) = \sum_{\omega\in\Omega} (X+Y)(\omega)P(\omega) = \sum_{\omega\in\Omega} X(\omega)P(\omega) + \sum_{\omega\in\Omega} Y(\omega)P(\omega) = E(X) + E(Y)$$

On calcule les espérances relativement aux événements élémentaires afin de pouvoir utiliser la linéarité de l'opérateur Σ.

En prenant Y constante égale à b, on obtient :

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

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

Exemple :

On lance 4 dés, et on note S la somme des résultats obtenus. Calculer $E(S)$.

Soient X_1, X_2, X_3 et X_4 les résultats obtenus pour chaque dé. On a :

$$E(X_1) = E(X_2) = E(X_3) = E(X_4) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = 3{,}5$$

Or, $S = X_1 + X_2 + X_3 + X_4$, d'où :

$$E(S) = E(X_1 + X_2 + X_3 + X_4) = 4E(X_1) = 4 \times 3{,}5 = 14$$

2.5. Théorème *Calcul pratique de la variance : formule de Koenig-Huyghens*

La variance d'une variable aléatoire X peut se calculer avec la relation suivante :

$$V(X) = E(X^2) - [E(X)]^2$$

la variance est l'écart entre la moyenne des carrés et le carré de la moyenne

Démonstration : on rappelle que l'espérance d'une variable aléatoire constante $X = b$ est égale à la constante b.

D'après la linéarité de l'espérance :

$$V(X) = E((X - E(X))^2) = E(X^2 - 2XE(X) + E(X)^2) = E(X^2) - 2E(X)E(X) + E(X)^2E(1)$$

$$V(X) = E(X^2) - [E(X)]^2$$

Pour le calcul de la variance, on préférera l'emploi de cette dernière formule plutôt que celle vue en 2.3. En effet, outre un intérêt pratique indéniable pour mener le calcul, la formule de Koenig-Huyghens est surtout plus fiable lorsque l'espérance $E(X)$ ne tombe pas juste. En effet, dans la formule vue en 2.3. l'erreur due à l'arrondi de $E(X)$ se propage tout au long du calcul alors qu'elle n'apparaît que dans le dernier terme dans la formule 2.5.

Exemple :

Reprenons l'exemple de la pièce de monnaie lancée trois fois de suite. X désigne le nombre de "face" obtenu.

$$E(X^2) = \frac{1}{8}\times 0^2 + \frac{3}{8}\times 1^2 + \frac{3}{8}\times 2^2 + \frac{1}{8}\times 3^2 = 3$$

$$V(X) = E(X^2) - [E(X)]^2 = 3 - \frac{9}{4} = \frac{3}{4}$$

2.6. Corollaire *Effet d'un changement affine sur la variance et l'écart type*

Soit X une variable aléatoire. Soient a et b deux réels.

On a : $$V(aX+b)=a^2V(X) \text{ et } \sigma(aX+b)=|a|\sigma(X)$$

En particulier : $$V(aX)=a^2V(X) \text{ et } \sigma(aX)=|a|\sigma(X)$$
$$V(X+b)=V(X) \text{ et } \sigma(X+b)=\sigma(X)$$

Démonstration :

D'après 2.5., on a : $$V(aX+b)=E(a^2X^2+2abX+b^2)-[E(aX+b)]^2$$

Et d'après la linéarité de l'espérance :

$$V(aX+b)=a^2E(X^2)+2abE(X)+b^2-[aE(X)+b]^2$$
$$V(aX+b)=a^2E(X^2)+2abE(X)+b^2-a^2[E(X)]^2-2abE(X)-b^2$$
$$V(aX+b)=a^2V(X)$$

D'où, par passage à la racine carrée : $$\sigma(aX+b)=|a|\sigma(X)$$

En particularisant $b=0$, puis $a=1$, on obtient :

$$V(aX)=a^2V(X) \text{ et } \sigma(aX)=|a|\sigma(X)$$
$$V(X+b)=V(X) \text{ et } \sigma(X+b)=\sigma(X)$$

Interprétation des dernières relations : une translation n'a pas d'incidence sur la variance ou l'écart type d'une variable aléatoire (en effet, cela ne modifie pas son degré de dispersion).

2.7. Définition *Fonction de répartition d'une variable aléatoire*

Soit X une variable aléatoire. La fonction de répartition F associée à X est la fonction définie sur $\mathbb{R}$ par :

$$F(x)=P(X\leqslant x)$$

La fonction de répartition est toujours une fonction croissante et bornée par 0 et 1.

Exemple : avec toujours les mêmes données précédentes, on a :

Pour $x\in\,]-\infty\,;0[$, on a : $$F(x)=0$$

Pour $x\in[0\,;1[$, on a : $$F(x)=\frac{1}{8}$$

Pour $x\in[1\,;2[$, on a : $$F(x)=\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$$

Pour $x\in[2\,;3[$, on a : $$F(x)=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}=\frac{7}{8}$$

Pour $x\in[3\,;+\infty[$, on a : $$F(x)=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=1$$

Représentation graphique :