

1 Espérance

1.1 Définition

En statistique, supposons que l'on relève une grandeur x . On peut supposer que l'on a différentes catégories d'individus que l'on peut étiqueter avec l'indice i . Une catégorie donnée aura une certaine fréquence f_i et une certaine valeur x_i .

On peut alors calculer la moyenne :

$$\bar{x} = \sum_i f_i \cdot x_i$$

Remarquez que deux i différents pourraient avoir le même x_i , cela n'a pas d'importance.

L'**espérance** est l'équivalent dans le domaine de probabilité. L'univers Ω se divise en issues (événements élémentaires) ω_i . Pour chaque issue, la variable aléatoire prend une valeur x_i et on a :

$$E(X) = \bar{X} = \sum_i x_i \cdot p(\omega_i)$$

Comme les probabilités correspondent aux fréquences obtenues si on répétait l'expérience un grand nombre de fois, alors l'espérance correspond à la moyenne que l'on obtiendrait si on répétait l'expérience aléatoire un grand nombre de fois.

On peut regrouper en semble les issues donnant un même x_i égal à k de sorte que la relation devient

$$E(X) = \bar{X} = \sum_k k \cdot p(X = k)$$

L'usage est d'utiliser un grand X pour indiquer que l'on parle d'une variable aléatoire.

1.2 Linéarité

L'espérance est linéaire, c'est à dire que

$$E(aX + bY) = a \cdot E(X) + b \cdot E(Y) \Rightarrow \overline{a \cdot X + b \cdot Y} = a \cdot \bar{X} + b \cdot \bar{Y}$$

Preuve : pour chaque issue ω_i , X prend une valeur x_i et Y prend une valeur y_i .

$$\begin{aligned} E(aX + bY) &= \sum_i (a \cdot x_i + b \cdot y_i) \cdot p(\omega_i) \\ &= a \cdot \sum_i x_i \cdot p(\omega_i) + b \sum_i y_i \cdot p(\omega_i) \\ &= a \cdot E(X) + b \cdot E(Y) \end{aligned}$$

Cette formule ne suppose rien sur les variables X et Y . Elle est juste même si elles ne sont pas indépendantes.

1.3 Cas de la loi binomiale



$$E(X) = n \cdot p$$

1.3.1 Première preuve simple

Soit X suivant une loi binomiale $\mathcal{B}(n; p)$.

On peut appeler X_1 le résultat au premier tirage. On dira $X_1 = 1$ en cas de succès à ce tirage et $X_1 = 0$ en cas d'échec.

On déduit donc que $E(X_1) = 0 \times (1 - p) + 1 \times p = p$.

On peut faire la même chose pour tous les X_i avec $1 \leq i \leq n$.

On peut dire que $X = X_1 + \dots + X_n$ et alors

$$\begin{aligned} E(X) &= E(X_1 + \dots + X_n) \\ &= E(X_1) + \dots + E(X_n) \\ &= p + \dots + p \\ &= n \cdot p \end{aligned}$$

1.3.2 Deuxième preuve plus calculatoire

$$\begin{aligned} E(X) &= \sum_k 0^n k \cdot p(X = k) \\ &= \sum_{k=0}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

Puisque le premier terme de la somme est pour $k = 0$ on peut le supprimer sans changer le résultat :

$$E(X) = \sum_{k=1}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

On peut ensuite poser $h = k - 1$, de sorte que quand $1 \leq k \leq n$, on aura $0 \leq h \leq n - 1$ et que tous les k sont remplacés par $h + 1$.

$$\begin{aligned} E(X) &= \sum_{h=0}^{n-1} (h+1) \cdot \binom{n}{h+1} p^{h+1} (1-p)^{n-(h+1)} \\ &= \sum_{h=0}^{n-1} (h+1) \cdot \frac{n!}{(h+1)!(n-h-1)!} p \cdot p^h (1-p)^{n-h-1} \end{aligned}$$

On peut dire que $n! = n \times (n-1)!$ et on a déjà écrit $p^{h+1} = p \cdot p^h$ ce qui nous permet de placer $n \cdot p$ en facteur de la somme.

On sait aussi que $(h+1)! = (h+1) \cdot h!$ ce qui permet de simplifier un facteur $(h+1)$.

$$\begin{aligned} E(X) &= n \cdot p \sum_{h=0}^{n-1} \frac{(n-1)!}{h!((n-1)-h)!} p^h (1-p)^{(n-1)-h} \\ &= n \cdot p \sum_{h=0}^{n-1} \binom{n-1}{h} p^h (1-p)^{n-h-1} \end{aligned}$$

Dans la somme précédente, on reconnaît des probabilités de forme $p(Y = h)$ avec Y suivant $\mathcal{B}(n-1; p)$. La somme de ces probabilités doit donc faire 1.

$$\begin{aligned} E(X) &= n \cdot p \sum_{h=0}^{n-1} \binom{n-1}{h} p^h (1-p)^{(n-1)-h} \\ &= n \cdot p \cdot 1 \\ &= n \cdot p \end{aligned}$$

2 Variance

2.1 Définition

En statistique, on observe que la moyenne du carré est toujours supérieure au carré de la moyenne :

$$\overline{x^2} \geq \bar{x}^2$$

On observe également que la différence entre ces quantités donne une idée de la façon dont les données sont dispersées.

$$V = \overline{x^2} - \bar{x}^2 \geq 0$$

Cette quantité est appelée variance. Comme dans la section précédente, on peut donner un équivalent probabiliste.

$$V(X) = E(X^2) - E(X)^2 = \overline{X^2} - \bar{X}^2 \geq 0$$

2.2 Espérance d'un produit

Considérons deux variables X et Y et l'espérance de leur produit $E(X \cdot Y)$.

X est susceptible de prendre des valeurs x_i et Y est susceptible de prendre des valeurs y_j . On aurait alors

$$E(X \cdot Y) = \sum_i \sum_j x_i \cdot y_j \cdot p(X = x_i \text{ ET } Y = y_j)$$

Dans le cas où X et Y sont **indépendantes**, on sait que (c'est la définition)

$$p(X = x_i \text{ ET } Y = y_j) = p(X = x_i) \cdot p(Y = y_j)$$

De sorte que dans ce cas, l'espérance devient

$$\begin{aligned} E(X \cdot Y) &= \sum_i \sum_j x_i \cdot y_j \cdot p(X = x_i) \cdot p(Y = y_j) \\ &= \sum_i x_i \cdot p(X = x_i) \cdot \sum_j y_j \cdot p(Y = y_j) \\ &= \sum_i x_i \cdot p(X = x_i) \cdot E(Y) \\ &= E(Y) \cdot \sum_i x_i \cdot p(X = x_i) \\ &= E(Y) \cdot E(X) \end{aligned}$$



Si X et Y indépendants, $E(X \cdot Y) = E(X) \cdot E(Y)$

2.3 Variance de la loi binomiale



$$V(X) = n \cdot p \cdot (1 - p)$$

On utilise peu cette formule. On utilise plutôt la formule de l'écart-type qui en est une conséquence.

2.4 Preuve

2.4.1 Preuve déjà assez compliquée

On sait déjà que $E(X) = n \cdot p$ dont $E(X)^2 = (n \cdot p)^2 = n^2 \cdot p^2$.

Nous allons maintenant déterminer l'expression de $E(X^2)$.

On utilise encore les variables X_i élémentaires comme pour l'espérance.

$$\begin{aligned} E(X^2) &= E((X_1 + \dots + X_n)^2) \\ &= E\left(\sum_{i,j} X_i \cdot X_j (X_1 + \dots + X_n)^2\right) = \sum_{i,j} E(X_i \cdot X_j) \end{aligned}$$

Comme on doit considérer tous les couples (i, j) , il y a n^2 termes dans la somme, mais ils ne fonctionnent pas tous de la même façon :

- On a d'abord les couples X_i et x_j pour lesquels $i \neq j$. Dans ce cas là, X_i et X_j sont indépendantes (par définition de la loi binomiale), et donc

$$E(X_i \cdot X_j) = E(X_i) \cdot E(X_j) = p \cdot p = p^2$$

- On a ensuite les cas où $i = j$. On ne peut pas faire le calcul précédent car X_i n'est pas indépendant de lui-même.

Mais c'est plus simple : X_i vaut 0 ou 1, donc on a toujours $X_i^2 = X_i$ de sorte que

$$E(X_i \cdot X_i) = E(X_i^2) = E(X_i) = p$$

Dans la somme $\sum_{i,j} E(X_i \cdot X_j)$, on aura n^2 couples (i, j) , dont n couples pour lesquelles $i = j$, ce qui laisse $n^2 - n$ couples pour lesquelles $i \neq j$.

$$E(X^2) = n \cdot p + (n^2 - n)p^2$$

On peut conclure

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= (n \cdot p + (n^2 - n)p^2) - n^2 \cdot p^2 \\ &= n \cdot p - n \cdot p^2 \\ &= n \cdot p \cdot (1 - p) \end{aligned}$$

2.4.2 Preuve si on aime beaucoup le calcul...

On se limitera à retrouver l'expression de $E(X^2)$.

$$\begin{aligned}
E(X^2) &= \sum_{k=0}^n k^2 \cdot p(X = k) \\
&= \sum_{k=0}^n k^2 \cdot \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^n k^2 \cdot \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{h=0}^{n-1} (h+1)^2 \cdot \binom{n}{h+1} p^{h+1} (1-p)^{n-h-1} \\
&= \sum_{h=0}^{n-1} (h+1)^2 \cdot \frac{n \cdot (n-1)!}{(h+1) \cdot h! \cdot (n-h-1)!} p^{h+1} (1-p)^{n-h-1} \\
&= n \cdot p \cdot \sum_{h=0}^{n-1} (h+1) \cdot \frac{(n-1)!}{h! \cdot (n-h-1)!} p^h (1-p)^{n-h-1} \\
&= n \cdot p \cdot \left[\sum_{h=0}^{n-1} h \cdot \frac{(n-1)!}{h! \cdot (n-h-1)!} p^h (1-p)^{n-h-1} + \sum_{h=0}^{n-1} \frac{(n-1)!}{h! \cdot (n-h-1)!} p^h (1-p)^{n-h-1} \right]
\end{aligned}$$

La deuxième somme est la somme des probabilités d'une variable Y suivant la loi $\mathcal{B}(n-1; p)$ et la première somme est l'espérance de cette même variable. La deuxième somme vaut donc 1 et la première vaut $(n-1) \cdot p$.

$$E(X^2) = n \cdot p [(n-1) \cdot p + 1] = n \cdot p + (n^2 - n) \cdot p$$

Ce qui est bien la même formule que celle trouvée dans la section précédente.

3 Écart-type

3.1 Définition

Comme on l'a dit, la variance est toujours positive : $V \geq 0$.

De plus, la variance est homogène à une moyenne de x^2 . Elle est donc dans l'unité de x^2 .

Il est donc pertinent de prendre la **racine carrée** de V , ce qui nous donne une grandeur bien définie (V est positif) et dans la même unité que x .

$$\sigma = \sqrt{V}$$

Ce que l'on peut faire également en probabilités.

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

3.2 Cas de la loi binomiale



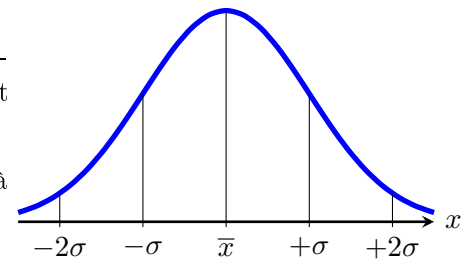
$$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)}$$

Étant donnée la formule de $V(X)$ trouvée précédemment, ce résultat est évident.

3.3 Intérêt de σ dans ce cas

L'écart-type est un très bon indicateur de la façon dont les données sont dispersées dans le cas particulier où les données sont réparties selon une **loi normale**.

Pour l'instant, il suffit de savoir que la loi normale correspond à une répartition selon la forme ci-contre.



Il se trouve que cette répartition est extrêmement courante dans tous les domaines des sciences et dans la nature, ce qui donne à σ un intérêt particulier.

De plus, la loi binomiale ressemble beaucoup à la loi normale. C'est d'ailleurs en observant les propriétés de la loi binomiale que l'on a découvert la loi normale.

Il sera donc très utile de savoir calculer $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$.