

1 Les notations

- X est une variable aléatoire. Quand on écrit X on parle d'une quantité qui n'a pas été encore mesurée. On fait des prévisions sur la valeur qu'aura cette quantité quand on l'aura mesurée.
- Comme nous parlons de variables aléatoires continues, on dit que la loi de probabilité de X a une densité que l'on peut noter δ .
- x est un nombre. Si j'écris par exemple $X = x$, c'est que j'envisage le cas où le résultat de ma mesure que je n'ai pas encore faite pourrait être égale à une certaine valeur x .
- δ est une fonction dont l'antécédent est une valeur de x . On peut donc écrire $\delta(x)$.

$\delta(x)$ exprime **en gros** à quel point il est probable que $X = x$.

Soyons plus précis : on a déjà dit qu'avec les lois continues, la probabilité que X ait une valeur exacte est 0. Autrement dit $p(X = x) = 0$.

Alors on aura plutôt $p(x \leq X < x + dx) = \delta(x) \cdot dx$, où dx est une quantité infinitésimale.

- y est une autre grandeur que l'on est susceptible de mesurer et elle dépend de x par une certaine fonction : $y = f(x)$. On ne veut pas alourdir les notations en ajoutant explicitement la fonction f alors on écrit y et sous-entendu, il s'agit de $y(x)$ de sorte que l'on pourra avoir aussi $y' = \frac{dy}{dx}$.
- Puisque x est aléatoire et prédit par X , y l'est aussi et on note Y la variable aléatoire nous permettant de faire des prédictions sur y . La fonction $x \mapsto y$ est la même que $X \mapsto Y$.

2 Densité et calcul de probabilité

2.1 Définition

Je vous rappelle la formule :

$$p(a \leq X < b) = \int_a^b \delta(x) \cdot dx$$

C'est le rôle de la fonction δ . Pour que cela fonctionne, il faut que :

- On puisse calculer l'intégrale, donc que δ ne soit pas une fonction trop bizarre.
- $\delta \geq 0$
- La probabilité pour toutes valeurs doit être de 1 donc

$$\int_{\text{toutes valeurs de } x} \delta(x) \cdot dx = 1$$

Important : Pour la suite je ne le préciserai plus, toutes les intégrales auront des bornes couvrant la totalité des valeurs de x . Par exemple, si je sais que x est forcément entre 0 et 10, on peut se contenter d'intégrer de 0 à 10. Si x peut prendre n'importe quelle valeur réelle, on doit intégrer de $-\infty$ à $+\infty$.

2.2 Cas d'une variable totalement dépendante

Prenons le cas maintenant de Y . Cette variable a sa propre densité que l'on peut noter Δ . Quel rapport entre Δ et δ dans l'hypothèse où Y dépend de X ?

$$p(x \leq X < x + dx) = \delta(x) \cdot dx$$

Si on arrondit au premier ordre, en supposant y localement croissante (donc $y'(x) > 0$:

$$x \leq X < x + dx \Leftrightarrow y(x) \leq Y(X) < y(x + dx) \Leftrightarrow y \leq Y < y(x) + \frac{dy}{dx} \cdot dx \Leftrightarrow y \leq Y < y + dy$$

On en déduit que $p(y \leq Y < y + dy) = p(x \leq X < x + dx)$ et donc :

$$\Delta(y) \cdot dy = \delta(x) \cdot dx \Leftrightarrow \Delta(y) \cdot \frac{dy}{dx} \cdot dx = \delta(x) \cdot dx$$

Quand on aura besoin de faire une intégrale sur y en faisant $\int \dots \Delta(y) \cdot dy$ on pourra donc remplacer :

$$\int \dots \Delta(y) \cdot dy = \int \dots \Delta(y) \cdot \frac{dy}{dx} \cdot dx = \int \dots \delta(x) \cdot dx$$

Si on fait la même chose dans le cas décroissant, on voit que cela fonctionne de la même façon.

2.3 Cas de deux variables indépendantes

Supposons que l'on ait deux variables X_1 et X_2 et que l'on veuille calculer la probabilité :

$$p(a \leq X_1 < b \quad \text{ET} \quad c \leq X_2 < d)$$

Il peut y avoir des liens entre X_1 et X_2 de sorte qu'en général il faudrait une densité commune $\delta_{1,2}(x_1, x_2)$ et le calcul de la proba serait :

$$p(a \leq X_1 < b \quad \text{ET} \quad c \leq X_2 < d) = \int_a^b \int_c^d \delta_{1,2}(x_1, x_2) \cdot dx_1 \cdot dx_2$$

Quand deux événements sont indépendants, on a $p(A \cap B) = p(A) \cdot p(B)$. De la même façon, on dira que X_1 et X_2 sont indépendantes si et seulement si $\delta_{1,2}(x_1, x_2) = \delta_1(x_1) \cdot \delta_2(x_2)$, c'est à dire si on peut séparer la densité de l'ensemble en un produit de deux densité indépendantes.

$$p(a \leq X_1 < b \quad \text{ET} \quad c \leq X_2 < d) = \int_a^b \int_c^d \delta_1(x_1) \cdot \delta_2(x_2) \cdot dx_1 \cdot dx_2 = \left(\int_a^b \delta_1(x_1) \cdot dx_1 \right) \cdot \left(\int_c^d \delta_2(x_2) \cdot dx_2 \right)$$

De sorte que l'on retrouve bien l'indépendance habituelle :

$$p(a \leq X_1 < b \cap c \leq X_2 < d) = p(a \leq X_1 < b) \cdot p(c \leq X_2 < d)$$

3 Espérance

3.1 Définition

L'espérance est une moyenne. Quand on calcule une moyenne, on fait le cumul pour toutes les valeurs possibles avec la pondération.

Ainsi, comme $p(x \leq X < x + dx) = \delta(x) \cdot dx$, la valeur x va compter avec une pondération $\delta(x) \cdot dx$ dans la moyenne.

Une somme dans un cas continu est une intégrale. On a donc :

$$E(X) = \bar{X} = \int x \cdot \delta(x) \cdot dx$$

3.2 Car Y dépendant de X

La formule est la même puisque l'on doit pouvoir traiter le cas Y sans savoir qu'il est dépendant de X . On a donc :

$$E(Y) = \bar{Y} = \int y \cdot \Delta(y) \cdot dy$$

Mais tenant compte de ce qu'on a dit plus haut :

$$E(Y) = \bar{Y} = \int y(x) \cdot \delta(x) \cdot dx$$

En mettant en évidence que y est fonction de x .

3.3 Espérance d'une constante

L'espérance étant une moyenne, on devine que si $X = K$ est constant, sa moyenne sera égale à la constante et donc $E(X) = K$. Cela se retrouve par le calcul :

$$E(K) = \int K \cdot \delta(x) \cdot dx = K \cdot \int \delta(x) \cdot dx = K \cdot 1 = K$$

3.4 La linéarité de l'espérance

On dit qu'une fonction F est linéaire quand on a $F(a \cdot X + b \cdot Y) = a \cdot F(X) + b \cdot F(Y)$. C'est donc le cas pour l'espérance.

Multiplication par une constante :

$$E(a \cdot X) = \int a \cdot x \cdot \delta(x) \cdot dx = a \cdot \int x \cdot \delta(x) \cdot dx = a \cdot E(X) \Leftrightarrow \overline{a \cdot X} = a \cdot \bar{X}$$

Somme :

$$E(Z + Y) = \int (z + y) \cdot \delta(x) \cdot dx = \int z \cdot \delta(x) \cdot dx + \int y \cdot \delta(x) \cdot dx = E(Z) + E(Y)$$

Ce que l'on peut résumer par $\overline{Z + Y} = \bar{Z} + \bar{Y}$.

Ces propriétés viennent du fait que l'espérance est une intégrale et que l'intégrale est linéaire.

On en déduit :

$$E(a \cdot X + b) = E(a \cdot X) + E(b) = a \cdot E(X) + b$$

Soit encore $\overline{a \cdot X + b} = a \cdot \bar{X} + b$.

3.5 Produit de variables indépendantes

$$\begin{aligned} E(X_1 \cdot X_2) &= \int \int x_1 \cdot x_2 \cdot \delta_1(x_1) \cdot \delta_2(x_2) \cdot dx_1 \cdot dx_2 \\ &= \left(\int x_1 \cdot \delta_1(x_1) dx_1 \right) \cdot \left(\int x_2 \cdot \delta_2(x_2) \cdot dx_2 \right) \\ &= E(X_1) \cdot E(X_2) \end{aligned}$$

Dit autrement, si X_1 et X_2 indépendantes :

$$\overline{X_1 \cdot X_2} = \bar{X}_1 \cdot \bar{X}_2$$

3.6 Espérance d'un carré

Peut-on simplifier $\overline{X^2}$? Est-ce égal à \bar{X}^2 ?

Non en général. X n'est pas indépendante d'elle-même, la formule de la section précédente ne s'applique donc pas.

Il n'y a donc pas de formule pour $\overline{X^2}$.

4 La variance

4.1 Définition

Je rappelle que par définition

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

On sait que cette quantité est toujours positive et on peut donc définir :

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

4.2 Cas $Y = a \cdot X + b$

Tout ce que nous avons dit plus haut va nous faire gagner du temps.

$$\begin{aligned} V(Y) &= \overline{Y^2} - \overline{Y}^2 \\ &= \overline{(a \cdot X + b)^2} - (\overline{a \cdot X + b})^2 \\ &= \overline{a^2 \cdot X^2 + 2 \cdot a \cdot b \cdot X + b^2} - (a \cdot \overline{X} + b)^2 \\ &= a^2 \cdot \overline{X^2} + 2 \cdot a \cdot b \cdot \overline{X} + b^2 - (a^2 \cdot \overline{X}^2 + 2 \cdot a \cdot b \cdot \overline{X} + b^2) \\ &= a^2 \cdot \overline{X^2} - a^2 \cdot \overline{X}^2 \\ &= a^2 \cdot (\overline{X^2} - \overline{X}^2) \\ &= a^2 \cdot V(X) \end{aligned}$$

Et on en déduit $\sigma(Y) = \sqrt{a^2} \cdot \sigma(X)$.

4.3 Cas $Y = a \cdot X_1 + b \cdot X_2 + c$

On peut faire le même travail. Je vais séparer les calculs de $\overline{Y^2}$ et \overline{Y}^2 , vous devriez voir comment cela fonctionne.

$$\begin{aligned} \overline{Y^2} &= \overline{(a \cdot X_1 + b \cdot X_2 + c)^2} \\ &= \overline{a^2 \cdot X_1^2 + b^2 \cdot X_2^2 + c^2 + 2ab \cdot X_1 X_2 + 2ac \cdot X_1 + 2bc \cdot X_2} \\ &= a^2 \cdot \overline{X_1^2} + b^2 \cdot \overline{X_2^2} + c^2 + 2ab \cdot \overline{X_1 \cdot X_2} + 2ac \cdot \overline{X_1} + 2bc \cdot \overline{X_2} \end{aligned}$$

$$\begin{aligned} \overline{Y}^2 &= \left(\overline{a \cdot X_1 + b \cdot X_2 + c} \right)^2 \\ &= (a \cdot \overline{X_1} + b \cdot \overline{X_2} + c)^2 \\ &= a^2 \cdot \overline{X_1}^2 + b^2 \cdot \overline{X_2}^2 + c^2 + 2ab \cdot \overline{X_1} \cdot \overline{X_2} + 2ac \cdot \overline{X_1} + 2bc \cdot \overline{X_2} \end{aligned}$$

On voit qu'une bonne part des termes se ressemblent. Quand on va faire la différence, seules les termes au carré vont rester :

$$\overline{Y^2} - \overline{Y}^2 = a^2 (\overline{X_1^2} - \overline{X_1}^2) + b^2 (\overline{X_2^2} - \overline{X_2}^2)$$

Et on trouve $V(Y) = a^2 \cdot V(X_1) + b^2 \cdot V(X_2)$.

On peut se convaincre que cela se généralise s'il y a plus de deux variables.

5 Cadre du TD : cas une variable

Dans le cas du TD on avait $y = f(x)$ mais avec l'hypothèse que x variait assez peu autour de sa moyenne \bar{X} de sorte que l'on pouvait approximer la fonction f au premier ordre par une fonction affine : $f(x) \approx a \cdot x + b$.

Puisqu'il est clair que c'est une approximation au premier ordre, j'écrirais = par la suite : c'est égal au premier ordre.

Donc dans le cadre de cette approximation, $y = a \cdot x + b$. On sait par ailleurs que $a = f'(\bar{X})$.

La suite devient évidente :

$$\bar{Y} = \overline{a \cdot X + b} = a \cdot \bar{X} + b = f(\bar{X})$$

C'est bien ce que l'on a observé pendant le TD.

$$V(Y) = V(a \cdot X + b) = a^2 V(X) \Rightarrow \sigma(Y) = \sqrt{a^2} \cdot \sigma(X)$$

Où $a = f'(\bar{X})$. Là encore, c'est bien ce que l'on a trouvé.

6 Cadre du TD : cas deux variables

Dans la 2e partie, on avait $y = f(x_1, x_2)$, avec encore x_1 et x_2 variant peu autour de leurs moyennes \bar{X}_1 et \bar{X}_2 .

On peut là encore faire une approximation au premier ordre :

$$y = a \cdot x_1 + b \cdot x_2 + c$$

avec $a = \left. \frac{\partial f}{\partial x_1} \right|_{x_1=\bar{X}_1, x_2=\bar{X}_2}$ et $b = \left. \frac{\partial f}{\partial x_2} \right|_{x_1=\bar{X}_1, x_2=\bar{X}_2}$.

Le principe est le même :

$$\bar{Y} = \overline{a \cdot X_1 + b \cdot X_2 + c} = a \cdot \bar{X}_1 + b \cdot \bar{X}_2 + c = f(\bar{X}_1, \bar{X}_2)$$

Et pour la variance

$$V(Y) = V(a \cdot X_1 + b \cdot X_2 + c) = a^2 \cdot V(X_1) + b^2 \cdot V(X_2)$$

Soit, avec les écarts-types :

$$\sigma(Y) = \sqrt{(a \cdot \sigma_{X_1})^2 + (b \cdot \sigma_{X_2})^2}$$