Probabilités et Statistiques

Pr. Hamza El Mahjour

Tanger, March 17, 2025

Outline

Espérance, variance et écart-type

Espérance

L'espérance d'une v.a X discrète s'exprime également par la formule

$$\mathbb{E}[X] = \sum_{s} s \times \mathbb{P}(\{X = s\})$$

Cette formule montre que l'espérance de X **ne dépend que de la loi de** X.

Cas d'une v.a.r continue

D'abord on définit la notion d'une v.a.r à densité f.

Definition

Une v.a.r X est dite à densité s'il existe une fonction f positive et intégrable sur $\mathbb{R}$ appelée **fonction de densité**, telle que pour tout a, b de $\mathbb{R}$ on ait

$$\mathbb{P}(\{a \leq X \leq b\}) = \int_a^b f(t)dt.$$

Donc la fonction de répartition dans le cas constinue est définie

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(u)du$$

F va être aussi croissane mais de plus elle est continue.

Dans le cas d'une variable aléatoire continue à densité [1] réelle, la définition précédente est adaptée comme suit

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \times f(x) dx$$

[1] La définition générale fait appel à la théorie de la mesure.

En fait, dans les définitions ci-dessus, on ne se préocuupe pas de l'existence des sommes $\sum_{\omega\in\Omega}$ ou $\sum_{s\in S}$ car on traite généralement des v.a.r avec un nombre fini de valeurs.

Alert Block

En revanche, si Ω est infini (dénombrable) il faut assurer que les sommes que l'on manipule sont bien définies (**ce n'est pas toujours le cas**).

v.a.r avec une infinité de valeurs

X prend toutes les valeurs de $\mathbb{N}^*$. Sa loi est définie par

$$\mathbb{P}(\{X=n\}) = \frac{1}{n(n+1)}$$

$\displaystyle\sum_n \mathbb{P}(\{X=n\})$ est bien une série convergente. Par contre l'espérance

$\displaystyle\mathbb{E}[X] = \sum_n n\mathbb{P}(\{X=n\}) = \sum_n \frac{1}{n+1}$ qui est une série divergente.

Espérance d'une fonction de X

Lorsque h est une fonction à valeurs réelles. L'espérance de $h(X)$ [si elle est définie] peut se mettre dans le cas discret sous la forme:

$$\mathbb{E}[h(X)] = \sum_{v \in S_{h(X)}} v\mathbb{P}(\{h(X) = v\}) = \sum_{s \in S_X} h(s)\mathbb{P}(\{X = s\}).$$

Dans le cas continu

$$\mathbb{E}[h(X)] = \int_{-\infty}^{+\infty} = h(u)f(u)du$$

absurdité...

N'écrivez pas $\mathbb{E}[h(X)] = \sum_{s \in S_X} h(s)h(\mathbb{P}(\{X = s\}))$ ou encore $\mathbb{E}[h(X)] = \int_{-\infty}^{+\infty} h(u)h(f(u))du$ ... fautes fréquentes chez les étudiants !

Outline

Variance et écart-type

Variance

La variance d'une v.a X mesure la déviance par rapport à la valeur moyenne ou à l'espérance.
Plus formellement, la variance est définie comme **l'espérance des écarts quadratiques de la variable et son espérance**

$$\mathbf{var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - \mathbb{E}[X]^2,$$

lorsque les variables X et $X - \mathbb{E}[X]$ possèdent une espérance.
Dans le cas continu

$$\mathbf{var}(X) = \int_{-\infty}^{+\infty} (s - \mathbb{E}[X])^2 f(s) ds.$$

Variance et écart-type

écart-type

On introduit aussi l'écart-type qui est définie par $\sigma(X) = \sqrt{\mathbf{var}(X)}$.
En fait, cette mesure permet de remettre les valeurs calculées dans l'ordre de l'unité de X.
Imaginez que la v.a.r X concerne des longueurs en m alors la variance sera en m^2 or l'écart-type sera avec la même unité m.
On peut dire que l'écart-type est un **paramètre d'échelle**.

Outline

Lois de variables aléatoires

Lois discrètes classiques

Lois de variables aléatoires

Lois continues classiques

Loi uniforme

Quand on veut modéliser un choix uniforme sur des intervalles de $\mathbb{R}$. Soit $[a,b] \subset \mathbb{R}$, alors la densité d'' une variable aléatoire qui suit une loi uniforme sur $[a,b]$ est:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si} \quad x \in [a,b], \\ 0 & \text{sinon.} \end{cases}$$

Dans ce cas on obtient la fonction de répartition

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ \frac{x-a}{b-a} & \text{si} \quad a \leq x < b \\ 1 & \text{si} \quad x \geq b \end{cases}$$

Son espérance et sa variance sont:

$$\mathbb{E}[X] = \frac{a+b}{2} \qquad \text{et} \qquad \mathbf{var}(X) = \frac{(b-a)^2}{12}.$$

Loi exponentielle de paramètre λ

Cette loi modélise généralement l'intervalle de temps séparant deux occurences successives d'un processus de Poisson. Ainsi, la probabilité qu'il n'y ait aucune occurence dans un intervalle de temps de longueur t est $p_0(t) = e^{-\lambda\, t}$.
$\rightarrow$ durée de vie de la radioactivité.
$\rightarrow$ durée de vie d'un composant électronique.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si} \quad x \geq 0, \\ 0 & \text{si.} \quad x < 0 \end{cases} \implies F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si} \quad x \geq 0, \\ 0 & \text{si.} \quad x < 0 \end{cases}$$

Donc

$$\mathbb{E}[X] = \frac{1}{\lambda} \qquad \text{et} \qquad \mathbf{var}(X) = \frac{1}{\lambda^2}.$$

Absence de mémoire

Principe markovien $\rightarrow$ phénomène est interprété par "une absence de mémoire"

En effet, pour tout $s, t \in \mathbb{R}$:

$$\mathbb{P}(\{X > t + s\} | \{X > t\}) = \mathbb{P}(\{X > s\})$$

Comment trouver la fonction de répartition de cette fonction?

Loi normale (de Gauss)

La loi normale **centrée réduite** est définie par la densité de probabilité

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

centrée $\rightarrow \mathbb{E}[X] = 0$

réduite $\rightarrow \mathbf{var}(X) = 0$

Calcul ...

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

$$x, y \in [0, +\infty[\longrightarrow r \in [0, +\infty[\quad \text{et} \quad \theta \in [0, \pi/2[$$

$$G = \int_0^{+\infty} e^{-x^2} dx \qquad \text{et} \qquad H = \int_0^{+\infty} \int_0^{+\infty} e^{-(x^2+y^2)} dx dy.$$

Fubini ...

Calcul ...

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

$$x, y \in [0, +\infty[\longrightarrow r \in [0, +\infty[\quad \text{et} \quad \theta \in [0, \pi/2[$$

$$G = \int_0^{+\infty} e^{-x^2} dx \qquad \text{et} \qquad H = \int_0^{+\infty} \int_0^{+\infty} e^{-(x^2+y^2)} dx dy.$$

Fubini ...

$$\phi(x) = \int_{-\infty}^{x} \varphi(t) dt$$

Calcul ...

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

$$x, y \in [0, +\infty[\longrightarrow r \in [0, +\infty[\quad \text{et} \quad \theta \in [0, \pi/2[$$

$$G = \int_0^{+\infty} e^{-x^2} dx \qquad \text{et} \qquad H = \int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)} dx dy.$$

Fubini ...

$$\phi(x) = \int_{-\infty}^{x} \varphi(t) dt$$

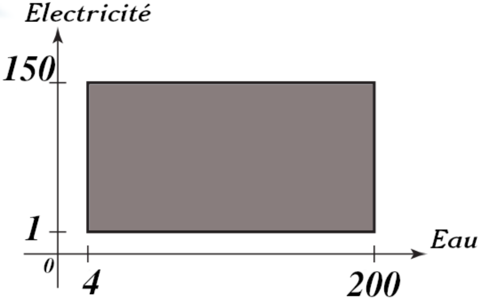
En pratique on utilisera un tableau !

Lois de variables aléatoires

Lois conjointes

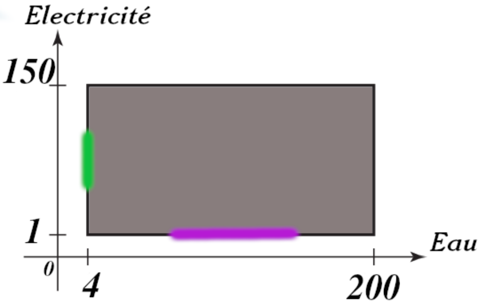
Exemple motivateur

Un entrepreneur planifie la construction d'un complexe de bureaux et doit estimer la demande en eau et en électricité. Il estime que la demande en électricité variera entre 1 et 150 millions de KWH par jour, tandis que la demande en eau sera comprise entre 4 et $200m^3$ par jour. Les combinaisons possibles des besoins de consommation en élécricité et en eau sont représentés sur la figure suivante: $X \to$ Eau, $Y \to$ Électricité.



Exemple motivateur

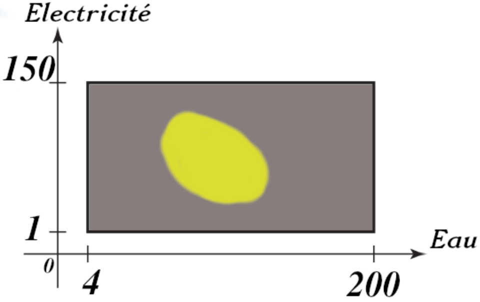
Un entrepreneur planifie la construction d'un complexe de bureaux et doit estimer la demande en eau et en électricité. Il estime que la demande en électricité variera entre 1 et 150 millions de KWH par jour, tandis que la demande en eau sera comprise entre 4 et $200 m^3$ par jour. Les combinaisons possibles des besoins de consommation en éléctricité et en eau sont représentés sur la figure suivante: $X \to$ Eau, $Y \to$ Électricité.



$\mathbb{P}(X \in \text{rose}) = \frac{\text{longueur en rose}}{200-4}$ $\qquad$ $\mathbb{P}(Y \in \text{vert}) = \frac{\text{longueur en vert}}{150-1}$

Exemple motivateur

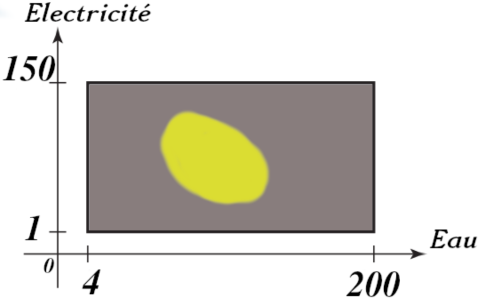
Un entrepreneur planifie la construction d'un complexe de bureaux et doit estimer la demande en eau et en électricité. Il estime que la demande en électricité variera entre 1 et 150 millions de KWH par jour, tandis que la demande en eau sera comprise entre 4 et $200 m^3$ par jour. Les combinaisons possibles des besoins de consommation en élécricité et en eau sont représentés sur la figure suivante: $X \to$ Eau, $Y \to$ Électricité.



$\mathbb{P}(X, Y) \in$ région jaune $=$

Exemple motivateur

Un entrepreneur planifie la construction d'un complexe de bureaux et doit estimer la demande en eau et en électricité. Il estime que la demande en électricité variera entre 1 et 150 millions de KWH par jour, tandis que la demande en eau sera comprise entre 4 et $200m^3$ par jour. Les combinaisons possibles des besoins de consommation en éléctricité et en eau sont représentés sur la figure suivante: $X \to$ Eau, $Y \to$ Électricité.



$$\mathbb{P}(X,Y) \in \text{région jaune} = \frac{\text{surface jaune}}{\text{surface grise}}$$

Definition

Loi conjointe Soit X et Y deux variables aléatoires. La **loi conjointe** de X et Y est la collection des probabilités de la forme $\mathbb{P}\left((X, Y) \in S\right)$ pour tout couple de valeurs réelles (x, y) appartenant à $S \in \mathbb{R}^2$

On va par la suite discuter comment caractériser et manipuler ces lois conjointes!

Lois conjointes discrètes

Definition

Fonction de distribution conjointe Si les v.a.r X et Y ont un nombre fini ou dénombrable de valeurs alors leur loi conjointe est discrète ainsi, on définit la fonction suivante:

$$f(x,y) = \mathbb{P}(X = x \text{ et } Y = y) = \mathbb{P}(X = x, Y = y)$$

Notons qu'on a le théorème suivant:

Theorem

Soit X, Y de loi conjointe discrète alors si (x_0, y_0) n'est pas une valeur possible de (X, Y) , alors $f(x,y) = 0$.
On a

$$\mathbb{P}\left((X,Y) \in S\right) = \sum_{(x,y)\in S} f(x,y).$$

De plus $\displaystyle\sum_{(x,y)\in\mathbb{R}^2} f(x,y) = 1.$

Tableau de loi conjointe

Voici un exemple pratique pour manipuler des lois conjointes. Dans un complexe résidentiel américain chaque maison a déclaré le nombre de voitures et de télévisions dont elles disposent. Soit X le nombre de voitures possédées par une personne choisie aléatoirement et Y le nombre de TVs.

On suppose que $X = 1, 2$ ou 3 et $Y = 1, 2, 3$ ou 4.

	y			
x	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

Par exemple, si on veut calculer $\mathbb{P}(X \geq 2, Y \geq 2)$ ça donne

$$\mathbb{P}(X \geq 2, Y \geq 2) = f(2,2) + f(2,3) + f(2,4) + f(3,2) + f(3,3) + f(3,4) = 0.5$$

Qu'en est-t-il de l'événement : "un foyer possède une seule voiture" ?

.

Tableau de loi conjointe

Voici un exemple pratique pour manipuler des lois conjointes. Dans un complexe résidentiel américain chaque maison a déclaré le nombre de voitures et de télévisions dont elles disposent. Soit X le nombre de voitures possédées par une personne choisie aléatoirement et Y le nombre de TVs.

On suppose que $X = 1, 2$ ou 3 et $Y = 1, 2, 3$ ou 4.

	y			
x	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

Par exemple, si on veut calculer $\mathbb{P}(X \geq 2, Y \geq 2)$ ça donne

$$\mathbb{P}(X \geq 2, Y \geq 2) = f(2,2) + f(2,3) + f(2,4) + f(3,2) + f(3,3) + f(3,4) = 0.5$$

Qu'en est-t-il de l'événement : "un foyer possède une seule voiture" ?

$$\mathbb{P}(X = 1) = \sum_{y=1}^{4} f(1, y) = 0.2 \, .$$

Tableau de loi conjointe

Voici un exemple pratique pour manipuler des lois conjointes. Dans un complexe résidentiel américain chaque maison a déclaré le nombre de voitures et de télévisions dont elles disposent. Soit X le nombre de voitures possédées par une personne choisie aléatoirement et Y le nombre de TVs.

On suppose que $X = 1, 2$ ou 3 et $Y = 1, 2, 3$ ou 4.

	y			
x	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

Par exemple, si on veut calculer $\mathbb{P}(X \geq 2, Y \geq 2)$ ça donne

$$\mathbb{P}(X \geq 2, Y \geq 2) = f(2,2) + f(2,3) + f(2,4) + f(3,2) + f(3,3) + f(3,4) = 0.5$$

Qu'en est-t-il de l'événement : "un foyer possède une seule voiture" ?

$\mathbb{P}(X = 1) = \sum_{y=1}^{4} f(1, y) = 0.2$. On parle de loi marginale !

Lois marginales

Si l'on s'intéresse à un événement sur X quelle que soit la valeur prise par Y, on obtient la loi de la v.a. X qui, dans le contexte d'un couple de v.a., est appelée **loi marginale**

$$\mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, Y \in \mathbb{R}).$$

De même pour la loi marginale de Y

$$\mathbb{P}(Y \leq y) = \mathbb{P}(X \in \mathbb{R}, Y \leq y).$$

Example

On peut compléter le tableau précédent pour obtenir ses lois marginales discrètes

	y				
x	1	2	3	4	
1	0.1	0	0.1	0	0.2
2	0.3	0	0.1	0.2	0.6
3	0	0.2	0	0	0.2
	0.4	0.2	0.2	0.2	1

Lois conjointes continues à densité

On revient à l'exemple des locaux avec un besoin en eau et électricité. Il est clair, que pour formaliser l'idée présentée de manière intuitive (fraction de l'aire de la région S sur la région grise totale) nous avons besoin de calculer la double intégrale suivante:

$$\mathbb{P}((X,Y)\in S)=\int_S\int\frac{1}{29904}dxdy$$

Definition

Deux variables X et Y continues disposent d'une loi/distribution conjointe s'il existe une fonction à positive f définie sur $\mathbb{R}\times\mathbb{R}$ telle que l'intégrale $\mathbb{P}((X,Y)\in S)=\int_S\int f(x,y)dxdy$ existe pour chaque portion S du plan. On dit alors que f est la densité de probabilité du couple de v.a.r (X,Y).

Exemple de densité conjointe continue

Dans l'exemple précédent la densité du couple est

$$f(x,y)=\begin{cases}1/29204 & \text{si} \quad (x,y)\in[4,200]\times[1,150],\\ 0 & \text{sinon.}\end{cases}$$

La fonction de densité conjointe doit vérifier l'égalité suivante pour être bien définie

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1.$$

Covariance

Soit X et Y deux v.a.r. On appelle **covariance** de X et Y l'expression

$$\mathbf{cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}[X])\cdot(Y-\mathbb{E}[Y])\right].$$

Pour une v.a.r discrète on a la formule

$$\mathbf{cov}(X,Y) = \sum_i \sum_j p_{ij}(x_i-\mathbb{E}[X])\cdot(y_j-\mathbb{E}[Y]).$$

On a quelques propriétés importantes à retenir

1. $\mathbf{cov}(X,Y) = \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y] = \sum_i \sum_j p_{ij}x_i y_j - \mathbb{E}[X]\mathbb{E}[Y]$.
2. Pour tout $\lambda \in \mathbb{R}$, $\mathbf{var}(X+\lambda Y) = \mathbf{var}(X) + 2\lambda\mathbf{cov}(X,Y) + \lambda^2\mathbf{var}(Y)$.
3. Si X et Y sont indépendantes alors $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ et $\mathbf{cov}(X,Y) = 0$ et $\mathbf{var}(X+Y) = \mathbf{var}(X) + \mathbf{var}(Y)$.

Corrélation/Causalité

La covariance est une porte pour mesurer le "degré de liaison linéaire" entre deux variables aléatoires ou deux tableaux de données statistiques. Voici un exemple troublant

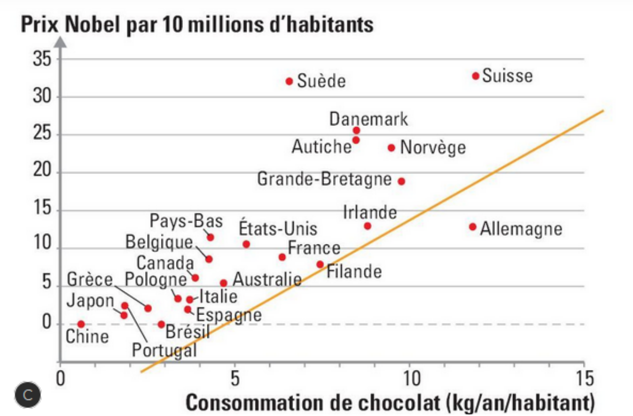


Figure: Consommons plus de chocolat! Gagnons le prix Nobel!

En fait, la corrélation est définie comme suit

$$r_{XY} = \frac{\mathbf{cov}(X,Y)}{\sigma_X \sigma_Y}.$$

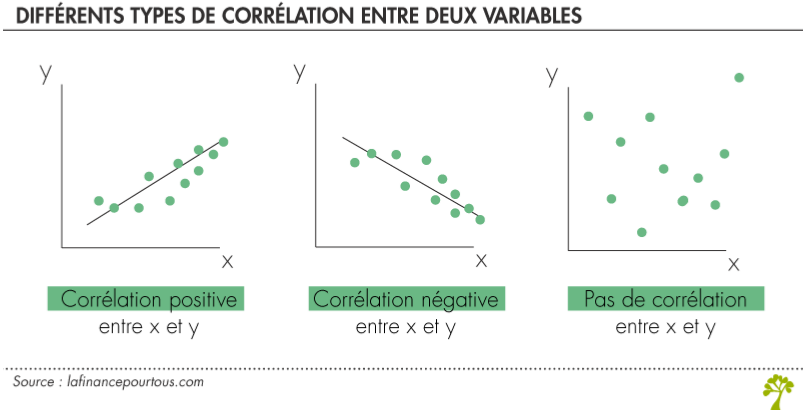
Propriétés

1. $|r_{XY}| \leq 1$.
2. Pour tout $a, b, c, d \in \mathbb{R}$ $r_{aX+b,cY+d} = r_{XY}$.
3. Si X et Y sont indépendants alors $r_{XY} = 0$.

Proof.

1. On sait que la variance est une quantité toujours positive, donc $\mathbf{var}(X+\lambda Y) \geq 0$ pour tout $\lambda \in \mathbb{R}$. Or $\mathbf{var}(X+\lambda Y) = \mathbf{var}(X) + 2\lambda\mathbf{cov}(X,Y) + \lambda^2\mathbf{var}(Y) := P(\lambda)$. Pour que le polynôme P soit toujours négatif il faut que le discriminant $\Delta_P = \mathbf{cov}(X,Y)^2 - \mathbf{var}(X)\mathbf{var}(Y) \leq 0$. Ainsi :
$r_{XY}^2 = \frac{\mathbf{cov}(X,Y)^2}{\mathbf{var}(X)\mathbf{var}(Y)} \leq 1 \implies |r_{XY}| \leq 1$. $\square$

Comment interpréter une corrélation?



1. r_{XY}: corrélation positive (forte $\simeq +1$)
2. r_{XY}: corrélation négative (forte : $\simeq -1$)
3. non corrélées linéairement : $r_{XY} \simeq 0$.

Une corrélation forte ne veut pas toujours dire une causalité.