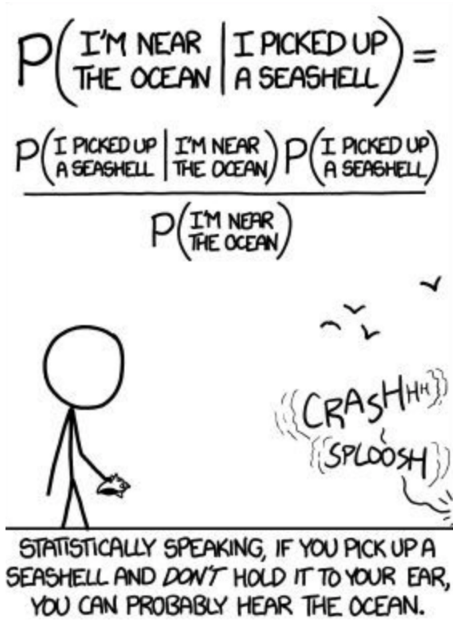
Probabilités : Loi de couples et inégalité de Bienaymé-Tchebychev



1 Loi d'un couple de variables aléatoires

Définition.

Étant données deux variables aléatoires discrètes X et Y, on appelle loi du couple de variables aléatoires (X, Y) la donnée des valeurs (x_i, y_j) avec leurs probabilités.

Exemple :

y_j \ x_i	-2	0	2
-1	0,1	0,1	0,2
2	0,2	0,2	0,2

En particulier, nous lisons dans le tableau que $P([X = 2] \cap [Y = -1]) = 0,2$.

Définition.

On appelle lois marginales les loi de X et de Y prises séparément. Elles apparaissent dans les marges du tableau comme sommes des lignes et des colonnes.

Exemple :

y_j \ x_i	-2	0	2	$P(Y = y_j)$
-1	0,1	0,1	0,2	0,4
2	0,2	0,2	0,2	0,6
$P(X = x_i)$	0,3	0,3	0,4	1

Voyons comment le tableau permet aussi de calculer la covariance de X et de Y :
$E(X) = -2 \times 0,3 + 0 \times 0,3 + 2 \times 0,4 = 0,2$ et $E(Y) = -1 \times 0,4 + 2 \times 0,6 = 0,8$
$E(XY) = 2 \times 0,1 - 2 \times 0,2 - 4 \times 0,2 + 4 \times 0,2 = -0,2$
Donc $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = -0,2 - 0,16 = -0,36$

Définition.

On appelle loi conditionnelle la loi d'une des deux variables lorsque l'autre a une valeur fixée.

Exemple :
On peut définir cinq lois conditionnelles à partir du précédent tableau :

x_i	-2	0	2
$P_{(Y=-1)}(X = x_i)$	0,25	0,25	0,5
$P_{(Y=2)}(X = x_i)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

y_j	$P_{(X=-2)}(Y = y_j)$	$P_{(X=0)}(Y = y_j)$	$P_{(X=2)}(Y = y_j)$
-1	$\frac{1}{3}$	$\frac{1}{3}$	0,5
2	$\frac{2}{3}$	$\frac{2}{3}$	0,5

Définition.

Le coefficient de corrélation linéaire de deux variables aléatoires X et Y est défini, à condition que $\sigma(X) \neq 0$ et $\sigma(Y) \neq 0$, par :

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

2 Inégalités de Markov et de Bienaymé-Tchebychev

Théorème : Inégalité de Markov.

Soit X une variable aléatoire, à valeurs positives, et qui admet une espérance.
Alors

$$\forall a > 0, \quad P(X \geq a) \leq \frac{E(X)}{a}$$

.

Démonstration.

Dans le cas discret :
$E(X) = \sum x_i P(X = x_i)$. On ne garde que les termes pour lesquels $x_i \geq a$.
Les autres termes étant tous positifs,
$E(X) \geq \sum_{\text{Les } x_i \text{ qui restent}} x_i P(X = x_i) \geq a \sum_{\text{Les } x_i \text{ qui restent}} P(X = x_i) = aP(X \geq a)$

Dans le cas continu :
$E(X) = \int_0^{+\infty} xf(x)dx \geq \int_a^{+\infty} xf(x)dx \geq a \int_a^{+\infty} f(x)dx = aP(X \geq a)$

Dans les deux cas on arrive à $E(X) \geq aP(X \geq a)$, c'est-à-dire $P(X \geq a) \leq \dfrac{E(X)}{a}$.

Théorème : Inégalité de Bienaymé-Tchebychev.

Soit X est une variable aléatoire qui admet une variance.
Alors

$$\forall \epsilon > 0, \quad P(|X - E(X)| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}$$

Démonstration.

Dans le cas discret :
$V(X) = \sum (x_i - E(X))^2 P(X = x_i)$
On ne garde que les termes pour lesquels $|x_i - E(X)| \geq \epsilon$.
Les autres termes étant tous positifs, $V(X) \geq \epsilon^2 \sum_{\text{Les } x_i \text{ qui restent}} P(X = x_i) = \epsilon^2 P(|X - E(X)| \geq \epsilon)$

Dans le cas continu :
$V(X) = \int_{-\infty}^{+\infty} (x-E(X))^2 f(x)dx \geq \int_{-\infty}^{E(X)-\epsilon} (x-E(X))^2 f(x)dx + \int_{E(X)+\epsilon}^{+\infty} (x-E(X))^2 f(x)dx$

donc $V(X) \geq \epsilon^2 \left(\int_{-\infty}^{E(X)-\epsilon} f(x)dx + \int_{E(X)+\epsilon}^{+\infty} f(x)dx \right) = \epsilon^2 P(|X - E(X)| \geq \epsilon)$.

Dans les deux cas on arrive à $V(X) \geq \epsilon^2 P(|X - E(X)| \geq \epsilon)$, c'est-à-dire $P(|X - E(X)| \geq \epsilon) \leq \dfrac{V(X)}{\epsilon^2}$.

3 Loi faible des grands nombres

Théorème.

Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de même espérance m et de même variance.
Pour tout n, on note $\overline{X_n}$ la variable aléatoire $\dfrac{X_1+\ldots+X_n}{n}$. Alors pour tout nombre ϵ strictement positif :

$$\lim_{n\to+\infty} P(|\overline{X_n}-m|\geq\epsilon)=0$$

Exemple.

On lance successivement un dé à six faces et pour tout $k\in\mathbb{N}^*$ on note X_k la variable aléatoire égale à la valeur obtenue au k$^{\text{ième}}$ lancer.
Les variables aléatoires $(X_k)_{k\in\mathbb{N}^*}$ sont indépendantes et pour tout $k\in\mathbb{N}^*$, $E(X_k)=3,5$ et $V(X_k)=\dfrac{35}{12}$.
Soit $\overline{X_n}=\dfrac{X_1+\ldots+X_n}{n}$ la moyenne des n premiers lancers.
D'après la loi faible des grands nombres, pour tout $\epsilon>0$, $\lim\limits_{n\to+\infty} P(|\overline{X_n}-3,5|\geq\epsilon)=0$.
Par exemple, on peut trouver un rang n suffisamment grand à partir duquel la probabilité que la moyenne des lancers soit comprise entre 3,49 et 3,51 soit supérieure à 0,99.