Probabilités A. El Ouni
Université Clermont Auvergne A. Khaldi
IUT Clermont Auvergne C. Samir
2A BUT Info, 2024-2025 A. Wohrer

TD8 – Plusieurs variables

Exercice 1.

1. On lance simultanément deux dés parfaits et indépendants. On introduit les deux variables aléatoires X=maximum des 2 dés, et Y=minimum des 2 dés.
   (a) Décrivez l'univers de cet expérience sous la forme d'un tableau carré. Indiquez la valeur de X et Y pour chacune des issues (=cases du tableau).
   (b) Calculez l'espérance de X et de Y.
   (c) Calculez la covariance de X et Y.
   (d) Calculez $P(X = 5|Y = 3)$.
   (e) Les variables X et Y sont elle indépendantes ?

Solution:

(a) TODO : tableau 6×6 (suivant les nombres sortis par chaque dé), dans lequel on rentre la valeur de X et Y dans chaque cas.

(b) $E(X) = 161/36$ et $E(Y) = 91/36$.

(c) On calcule d'abord $E(XY) = 12.25$. Ce calcul peut se faire de 2 manières :
- En notant la valeur de XY dans chacune des 36 cases du tableau, puis en calculant sa valeur moyenne.
- En remarquant (astuce !) qu'on a toujours $XY = D_1D_2$, or les variables D_1 et D_2 sont indépendantes donc $E(D_1D_2) = E(D_1)E(D_2) = 3.5^2 = 12.25$.

On a donc ensuite

$$Cov(X, Y) = E(XY) - E(X)E(Y) \simeq 0.95$$

(d) En se limitant aux cases du tableau pour lesquelles $Y = 3$:

$$P(X = 5|Y = 3) = P(X = 5 \text{ et } Y = 3)/P(Y = 3) = \frac{2}{7}$$

(e) Les variables X et Y ne sont pas indépendantes. Deux arguments possibles pour le prouver : leur covariance est non nulle. OU BIEN : $P(X = 5|Y = 3) \neq P(X = 5)$.

2. Mêmes questions pour les variables X=somme des deux dés et Y=écart (positif) entre les 2 dés.

Solution:

(a) TODO : tableau 6×6 (suivant les nombres sortis par chaque dé), dans lequel on rentre la valeur de X et Y dans chaque cas.

(b) $E(X) = 7$ et $E(Y) = 70/36$.

(c) En calculant la valeur de XY dans chacune des 36 cases du tableau, puis en calculant sa valeur moyenne, on obtient $E(XY) = 490/36$. On a donc ensuite

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{490}{36} - \frac{7.70}{36} = 0$$

(d) En se limitant aux cases du tableau pour lesquelles $Y = 3$:

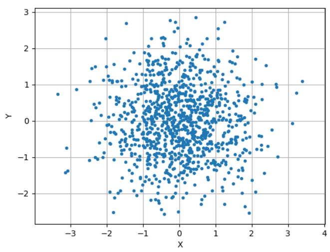
$$P(X = 5|Y = 3) = P(X = 5 \text{ et } Y = 3)/P(Y = 3) = \frac{2}{6} = \frac{1}{3}$$

(e) Les variables X et Y ne sont pas indépendantes, car $P(X = 5|Y = 3) \neq P(X = 5)$ (qui vaut $\frac{4}{36} = \frac{1}{4}$).
Remarque : pourtant on avait $Cov(X,Y) = 0$, mais cet argument ne suffit pas à garantir que des variables sont indépendantes.

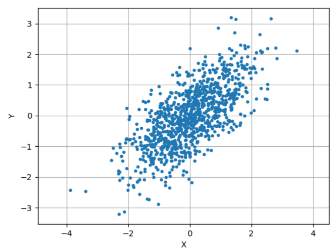
Exercice 2.

Chacune des figures ci-dessous représente un échantillon de tirages de 2 variables aléatoires (X,Y). Pour chacune des figures, répondez aux questions suivantes :

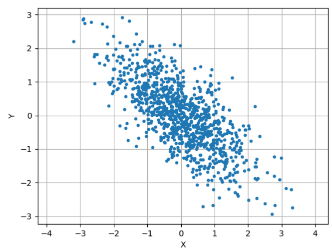
- La covariance de X et Y semble-t-elle positive, négative, ou nulle ?
- Les variables X et Y semblent-elles indépendantes ?



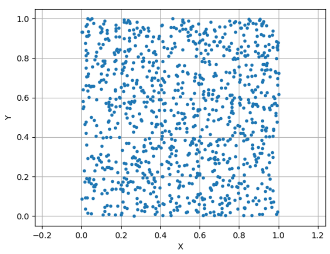
(a)



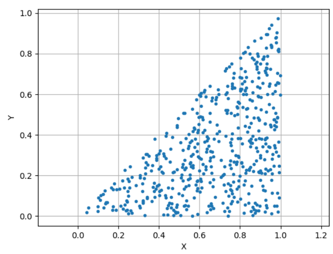
(b)



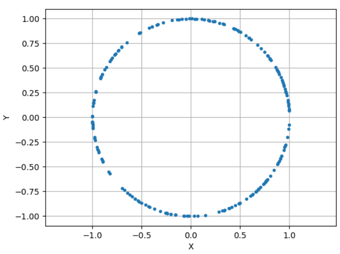
(c)



(d)



(e)



(f)

Solution:

(a) Covariance $\simeq 0$, et les variables semblent indépendantes. (La distribution de la variable Y ne semble pas dépendre de la vameur de X, et réciproquement.)

(b) Covariance > 0, les variables ne sont donc pas indépendantes.

(c) Covariance < 0, les variables ne sont donc pas indépendantes.

(d) Covariance $\simeq 0$, et les variables semblent indépendantes. (La distribution de la variable Y ne semble pas dépendre de la vameur de X, et réciproquement.)

(e) Covariance > 0 (la valeur moyenne de Y sachant X augmente lorsque X augmente), les variables ne sont donc pas indépendantes.

(f) Covariance $\simeq 0$. Pourtant les variables ne sont clairement PAS indépendantes. En effet, la valeur de X influe grandement sur la valeur de Y, puisqu'à X donné, Y peut uniquement prendre 2 valeurs imposées (demi-cercle du haut ou demi-cercle du bas).

Exercice 3.

Dans une colonie de marmottes de Sibérie orientale, on note la durée de vie X d'une marmotte au hasard. On suppose que X suit une loi exponentielle, de paramètre $\lambda = 0.5$ an^{-1}.

1. On définit Y la variable aléatoire donnant la moyenne des durées de vie de 3 marmottes prises au hasard. C'est-à-dire $Y = (X_1 + X_2 + X_3)/3$, avec X_1, X_2, X_3 trois variables aléatoires indépendantes suivant une loi exponentielle de paramètre λ. Donnez les valeurs de $E(Y)$ et $V(Y)$.

Solution: Déjà, d'après les propriétés de la distribution exponentielle, on a $E(X) = 1/\lambda = 2$ ans, et $V(X) = 1/\lambda^2{=}4$ ans^2.
Ensuite, par linéarité de l'espérance, $E(Y) = (E(X_1) + E(X_2) + E(X_3))/3 = (2 + 2 + 2)/3 = 2$ ans.
Ensuite, comme les variables X_1, X_2, X_3 sont **indépendantes**, un résultat de cours garantit que leurs variances s'aditionnent. On a donc

$$V(X_1 + X_2 + X_3) = V(X_1) + V(X_2) + V(X_3) = 4 + 4 + 4 = 12$$

Enfin, une variance représente une valeur moyenne du *carré* de la variable. Donc lorsqu'on multiplie la variable par une constante c, sa variance est multipliée par c^2. (Preuve : $V(cX) = E(c^2X^2) - E(cX)^2 = c^2(E(X^2) - E(X)^2) = c^2V(X)$.)
On en déduit que

$$V(Y) = V\big(\frac{X_1 + X_2 + X_3}{3}\big) = \frac{V(X_1 + X_2 + X_3)}{9} = \frac{12}{9} = \frac{4}{3}$$

2. On définit Z la variable aléatoire donnant la moyenne des durées de vie de 100 marmottes prises au hasard. C'est-à-dire $Z = (X_1 + \cdots + X_{100})/100$, avec $X_1, \ldots, X_{100}$ cent variables aléatoires indépendantes suivant une loi exponentielle de paramètre λ. Donnez les valeurs de $E(Z)$ et $V(Z)$.

Solution: Par le même argument mais en remplaçant 3 par 100, on a

$$E(Z) = 2$$

et

$$V(Z) = \frac{V(X)}{100} = \frac{4}{100} = 0.04$$

3. Donnez, en le justifiant, une bonne approximation pour la loi de la variable Z. Utilisez cette approximation pour calculer $P(1.8 \leq Z \leq 2.2)$.

Solution: Z est une somme d'un grand nombre de variables indépendantes. D'après le théorème central limite, la distribution de Z est donc approximativement une loi normale.
De plus, on a calculé au-dessus l'espérance (2) et l'écart-type ($\sqrt{0.04} = 0.2$) de Z. Donc

$$Z \sim \mathcal{N}(2, 0.2)$$

En particulier,

$$P(1.8 \leq Z \leq 2.2) = P(\frac{1.8 - 2}{0.2} \leq \frac{Z - 2}{0.2} \leq \frac{2.2 - 2}{0.2}) = P(-1 \leq T \leq 1)$$

pour T une loi normale centrée réduite. Il n'y a plus qu'à aller lire le tableau.

4. Soit un nombre entier N. On note W_N la variable aléatoire donnant la moyenne des durées de vie de N marmottes prises au hasard. Quelle valeur minimum doit avoir N pour qu'on ait $P(1.99 \leq W_N \leq 2.01) \geq 0.95$?

Solution: En répétant le même raisonnement, la variable W_N a toujours une espérance de 2, et une variance de

$$Var(W_N) = \frac{Var(X)}{N} = \frac{4}{N}$$

et de plus, dès que N est suffisamment grand, la distribution de W_N devient une loi normale. D'après le tableau pour la loi normale, le z correspondant à un intervalle symétrique à 95% est $z \simeq 1.96$. On doit donc résoudre

$$1.96 \times \sqrt{\frac{4}{N}} \leq 0.01$$

ce qui donne $N \geq 4 \times 196^2 \simeq 160000$.