IUT HSE — Probabilités et statistiques
UNIVERSITÉ DE BORDEAUX 1 — 2013/2014

Contrôle d'octobre 2014.
Documents interdits

On rappelle que si Z suit $\mathcal{N}(0,1)$ alors $P(|Z| > 1.96) \leq 0.05$ et $P(|Z| > 2.5758) \leq 0.01$.

Pour les applications numériques, on se contentera de $P(|Z| > 2) \leq 0.05$ et $P(|Z| > 2.6) \leq 0.01$.

Exercice 1.

(1) Calculer C_5^2 et C_8^7.
(2) De combien de façons peut-on classer (sans ex aequo) 27 étudiants d'un groupe de TD ?
 Peut-on les lister?

Solution de l'exercice 1. $C_5^2 = 10$ (triangle de Pascal ou formule du cours) et $C_8^7 = 8$ nombre de façons de choisir 7 personnes parmi 8: il suffit de savoir qui exclure, il y a 8 possibilités.

Il y a 27! façons de classer 27 personnes, ce nombre a (bien plus) de 20 chiffres, inutile de vouloir lister toues les possibilités.

Exercice 2.
En cas de survenue de maux de tête, 40% des patients prennent un médicament à base de paracétamol et 30% des patients prennent un médicament à base d'aspirine. Pour le premier, le taux de guérison en 24h est de 80% et pour le second de 90%. 20% des patients ne prennent aucun médicament dans les 24 premières heures et dans 60% des cas ils sont guéris le lendemain.

Enfin, les 10% restant prennent des médicaments prescrits sur ordonnance et dans 95% des cas, leurs maux de tête disparaissent en $24h$.

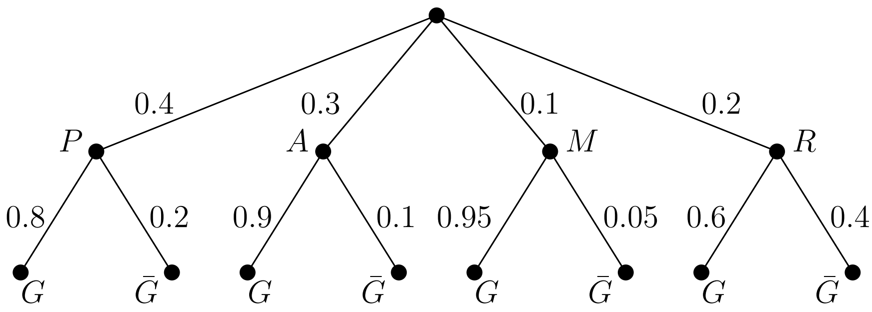
(i) Quel est le taux global de personnes guéries après $24\,h$?
(ii) Quel est la probabilité pour un patient de ne pas avoir pris de médicament sachant qu'il est guéri ?
(iii) Quel est la probabilité pour un patient d'avoir pris un médicament délivré sur ordonnance sachant qu'il est guéri ?

Solution de l'exercice 2. Il suffit de faire un arbre. On introduit

notation	événement
G	être guéri
A	prendre de l'aspirine
P	prendre du paracétamol
M	prendre un médicament sur ordonnance
R	ne rien prendre

et $\bar{X}$ est l'événement contraire de X.

Alors



Alors

$$\mathbb{P}(G) = 0.4 \times 0.8 + 0.3 \times 0.9 + 0.1 \times 0.95 + 0.2 \times 0.6 = 0.805$$

Puis

$$\mathbb{P}(R|G) = \frac{\mathbb{P}(R \cap G)}{\mathbb{P}(G)} = \frac{\mathbb{P}(G|R)\mathbb{P}(R)}{\mathbb{P}(G)} = \frac{0.6 \times 0.2}{0.805} = 0.15$$

et

$$\mathbb{P}(M|G) = \frac{\mathbb{P}(G|M)\mathbb{P}(M)}{\mathbb{P}(G)} = \frac{0.95 \times 0.1}{0.805} = 0.12$$

Exercice 3. Dans une population de canards colverts, lors d'une alerte, l'ensemble de la population quitte son aire de repos. Ainsi, à l'instant $t = 0$, l'aire est déserte. La probabilité pour qu'un canard revienne à l'étang est donnée par une densité de de la forme

$$f(t) = \begin{cases} 0 & \text{si } t < 0 \\ c\big(e^{-t} - e^{-at}\big) & \text{pour } t \geq 0 \end{cases}$$

avec $c > 0$ et $a > 1$.

(1) Montrez que $f \geq 0$ et déterminez c pour que $\displaystyle\int_{\mathbb{R}} f(t)\,\mathrm{d}t = 1$.

(2) On suppose maintenant que $a = 2$.

  (a) Calculez la probabilité pour qu'un canard revienne dans les 3 premières minutes.

  (b) Déterminez le temps moyen de retour.

Solution de l'exercice 3. si $a > 1$, et $t \geq 0$, $at \geq t$ donc $e^{-at} \leq e^{-t}$ puisque $x \mapsto e^{-x}$ est décroissante. Comme $c \geq 0$, $c(e^{-t} - e^{-at}) \geq 0$. Ensuite,

$$1 = \int_{\mathbb{R}} f(t)\,\mathrm{d}t = \int_0^{+\infty} c(e^{-t} - e^{-at})\,\mathrm{d}t = \left[c\left(-e^{-t} + \frac{e^{-at}}{a}\right)\right]_0^{+\infty} = -c\left(-1 + \frac{1}{a}\right)$$

puisque en $+\infty$, e^{-t} et e^{-at} tendent vers 0. On en déduit $c\frac{a-1}{a} = 1$ donc $c = \frac{a}{a-1}$.

En particulier, pour $a = 2$, $c = 2$. La probabilité pour qu'un canard revienne dans les 3 premières minutes est alors

$$\begin{aligned} \int_{-\infty}^{3} f(t)\,\mathrm{d}t &= \int_0^3 2(e^{-t} - e^{-2t})\,\mathrm{d}t = \left[2\left(-e^{-t} + \frac{e^{-2t}}{2}\right)\right]_0^3 \\ &= 2\left(-e^{-3} + \frac{e^{-6}}{2}\right) - 2\left(-e^{-0} + \frac{e^{-0}}{2}\right) = 0.9 \end{aligned}$$

et le temps moyen de retour est, après une intégration par parties,

$$\begin{aligned}
\int_{\mathbb{R}} tf(t)\,\mathrm{d}t &= \int_0^{+\infty} 2t(e^{-t} - e^{-2t})\,\mathrm{d}t \\
&= \left[2t(-e^{-t} + e^{-2t}/2)\right]_0^{+\infty} - \int_0^{+\infty} 2(-e^{-t} + e^{-2t}/2)\,\mathrm{d}t \\
&= 0 - 0 + \int_0^{+\infty} 2e^{-t} - e^{-2t}\,\mathrm{d}t \\
&= \left[-2e^{-t} + e^{-2t}/2\right]_0^{+\infty} = 0 - (-2 + 1/2) = 3/2
\end{aligned}$$

c'est-à-dire $1min30s$.

Exercice 4. On donne les effectifs par âge, de mères à l'accouchement dans la maternité de Bordeaux Nord la semaine dernière.

âge	21	22	23	24	25	26	27	28	29	30
effectif	0	2	5	6	7	9	10	5	2	4

(1) Calculer la moyenne $\tilde{\mu}$, la variance $\tilde{\sigma}^2$ et l'écart-type empiriques $\tilde{\sigma}$ de l'échantillon.
(2) Déterminez des intervalles de confiance pour l'estimation de l'âge moyen à 95% et à 99%.

Solution de l'exercice 4. Il y a $2 + 5 + \cdots + 4 = 50$ naissances.

$\tilde{\mu} = \dfrac{2 \times 22 + 5 \times 23 + \cdots + 4 \times 30}{50} = 26$ puis

$$\tilde{\sigma}^2 = \frac{2 \times 22^2 + 5 \times 23^2 + \cdots + 4 \times 30^2}{50} - 26^2 = 4.4$$

donc $\tilde{\sigma} = \sqrt{4.4} = 2.1$.

Exercice 5. Nous souhaitons exprimer la hauteur H d'un arbre en fonctions de son diamètre D à $1\,m\,30$ du sol. Pour cela, nous avons mesuré 50 couples diamètre-hauteur $(H_1, D_1), \ldots, (H_{50}, D_{50})$ et les résultats sont ci-dessous.

$$\mu_H = \frac{1}{50}\sum_{i=1}^{50} H_i = 35, \quad \frac{1}{50}\sum_{i=1}^{50}(H_i - \mu_H)^2 = 25; \mu_D = \frac{1}{50}\sum_{i=1}^{50} D_i = 2.5, \quad \frac{1}{50}\sum_{i=1}^{50}(D_i - \mu_D)^2 = 0.25$$

et $\dfrac{1}{50}\displaystyle\sum_{i=1}^{50}(H_i - \mu_H)(D_i - \mu_D) = 2.4$.

(1) Calculez le coefficient de corrélation. Que pouvez vous en conclure ?
(2) Donner l'équation de la droite de régression de H par rapport à D.
(3) Donner la hauteur prévue d'un arbre de diamètre $2\,m$.

Solution de l'exercice 5. $\sigma_H^2 = \dfrac{1}{50}\displaystyle\sum_{i=1}^{50}(H_i - \mu_H)^2 = 25$ donc $\sigma_H = 5$,

$\sigma_D^2 = \dfrac{1}{50}\displaystyle\sum_{i=1}^{50}(D_i - \mu_D)^2 = 0.25$ donc $\sigma_D = 0.5$

$\sigma_{H,D} = \frac{1}{50}\sum_{i=1}^{50}(H_i - \mu_H)(D_i - \mu_D) = 2.4.$ donc

$$\rho_{H,D} = \frac{\sigma_{H,D}}{\sigma_H \sigma_D} = \frac{2.4}{5 \times 0.5} = 0.96$$

qui est très près de 1. Il est donc raisonable de penser que le diamètre d'un arbre est linéairement relié à sa hauteur.

On cherche une relation de la forme $H = aD + b$, alors $a = \rho_{H,D}\frac{\sigma_H}{\sigma_D} = 9.6$ et $\mu_H = a\mu_D + b$ donc $b = 35 - 9.6 \times 2.5 = 11\,m$.

Pour $D = 2\,m$ ce modèle prédit $H = 9.6 \times 2 + 11 = 30.2\,m$.